DECONSTRUCTION AND RECONSTRUCTION

OF A PROTEIN CAPSID

by

Hsiao-Nung Chen

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Chemistry

The University of Utah

August 2014

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Hsiao-Nung Chen**

has been approved by the following supervisory committee members:

Kenneth J. Woycechowsky	, Chair	**4/10/2014** Date Approved
Peter F. Flynn	, Member	**4/10/2014** Date Approved
Jennifer M. Heemstra	, Member	**4/10/2014** Date Approved
Ilya Zharov	, Member	**4/10/2014** Date Approved
Edward M. Trujillo	, Member	**4/10/2014** Date Approved

and by **Cynthia J. Burrows** , Chair/Dean of

the Department/College/School of **Chemistry**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The self-assembly of multiple protein subunits via noncovalent interactions provides a diverse collection of nanoarchitectures. Polyhedral capsids represent a particularly interesting type of structure in which the protein forms a closed three-dimensional surface that can function as a molecular container. Understanding capsid self-assembly could benefit further development of nanomaterials for many applications, such as drug delivery and biocatalysis. In this thesis, *Aquifex aeolicus* lumazine synthase (AaLS) is used as a model for investigating capsid self-assembly. The 60-subunit capsids formed by AaLS *in vivo* can be viewed as dodecamers-of-pentamers. Currently, methods are lacking for controlling AaLS assembly *in vitro*, which imposes important limitations on cargo loading. Interestingly, the diverse quaternary structures in the lumazine synthase family, which includes capsids and pentamers, imply the possibility of exchanging assembly states *in vitro*. To better understand AaLS capsid assembly, the dodecahedral capsid was converted to pentamers via a strategy involving rational design and site-directed mutagenesis. Biophysical characterizations confirm that simultaneous substitution of three interfacial residues can yield stable pentamers. A pentameric AaLS variant that possesses a unique, strategically placed cysteine residue was engineered to enable capsid formation *in vitro*. This cysteine was modified with a thiophenol group by sequential thiol-disulfide exchange reactions. The increase in nonpolar surface area upon

modification of this cysteine allows for assembly of the pentamers into capsids that resemble wild-type AaLS, presumably by recapitulating hydrophobic interactions present at the pentamer-pentamer interface. In an alternative approach, a pH-dependent switch for AaLS capsid disassembly was developed. For this switch, the ability to change assembly state relies on the presence of three engineered histidine residues per subunit, located near the three-fold symmetric interface of the capsid. These histidines minimally interfere with the capsid structure at high pH where their side chains are neutral. However, at lower pH, histidine protonation can trigger the dissociation of the capsid into pentamers, presumably due to charge repulsion. Further, this switch is reversible, as 60-subunit capsids can reform upon raising the pH. These studies of interconverting AaLS quaternary structures open the door to the development of improved encapsulation systems for use in medicine or nanobiomaterials.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AaLS	*Aquifex aeolicus* lumazine synthase
AdhD	alcohol dehydrogenase D
AUC	analytical ultracentrifugation
β-ME	2-mercaptoethanol
BaLS	*Brucella abortus* lumazine synthase
BsLS	*Bacillus subtilis* lumazine synthase
CA	carbonic anhydrase
CD	circular dichroism
CCMV	cowpea chlorotic mottle virus
Da	dalton
DTT	dithiothreitol
DTNB	5,5′-dithiobis-(2-nitrobenzoate)
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
EDTA	ethylenediaminetetraacetic acid
ELPs	elastin-like polypeptides
ESI-MS	electrospray ionization mass spectrometry
FPLC	fast protein liquid chromatography

GdDOTA	gadolinium-tetraazacyclododecane tetraacetic acid
GFP	green fluorescent protein
HIV	human immunodeficiency virus
HRP	horseradish peroxidase
Hsp	heat shock protein
HNS	N-hydroxysuccinimide
IPTG	isopropyl β-D-1-thiogalactopyranoside
LS	lumazine synthase
MRI	magnetic resonance imaging
MPV	murine polyomavirus
MWCO	molecular weight cut off
NTB	2-nitro-5-thiobenzoate
PDB	protein data bank
PEG	polyethylene glycol
PET	positron emission tomography
RubisCO	ribulose bisphosphate carboxylase-oxygenase
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
SEC	size-exclusion chromatography
ScLS	*Saccharomyces cerevisiae* lumazine synthase
TEM	transmission electron microscopy

TECP	tris (2-carboxyethyl) phosphine
UV	ultraviolet

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Foremost, I would like to thank Dr. Kenneth Woycechowsky, my research advisor, for his mentorship, guidance, support, and patience for my research and study. I would also like to thank Dr. Jim Muller for assistance collecting and analyzing mass spectra, Dr. Debbie Eckert for assistance with the sedimentation equilibrium experiments, Mr. Michael Standing of the BYU microscopy lab for assistance collecting the EM images, and Dr. Mark Ji for assistance with SYBYL-X 2.0.

Thanks to all members of the Woycechowsky group for their helpful discussions and friendship. Special thanks to my wife, An-Chueh Chou, and my parents for their support and patience.

CHAPTER 1

INTRODUCTION

1.1. Protein Oligomerization

1.1.1. How Does Protein Oligomerization Arise?

Proteins often form oligomers from the association of multiple polypeptide subunits.[1,2] Indeed, a survey of protein oligomeric states in *Escherichia coli* reveals that approximately 75% of the proteins are oligomers, predominantly homooligomers (~79 % of the oligomeric proteins in *E. coli*).[1] Oligomeric proteins are essential elements for building many large architectures, such as cytoskeletal frameworks or protein capsids, which play important roles in many biological functions. In addition, proteins that form large architectures can also benefit from higher stability against degradation and denaturation.[3] Interestingly, although many biological functions demand large structures, natural evolution has usually selected noncovalent symmetric assemblies of small or medium sized protein subunits to build large architectures rather than long single polypeptide chains.

Several evolutionary advantages have been proposed for building large structure from repeating small subunits.[1] First, only minimal genome space is required for encoding a small subunit of a large assembly, which increases the coding efficiency compared to encoding a large polypeptide.[4] In addition, producing long peptide chains could increase the risks of translation errors during protein synthesis.[5,6] Further, building a large structure from many small subunits can decrease the biosynthetic cost required to fix translation errors because longer polypeptides are more expensive to discard and resynthesize than shorter ones.

Oligomeric proteins predominantly form symmetric structures by using identical subunits.[1,7–9] To explain the great abundance of symmetric homooligomers occurring in

Nature, many advantages linked to symmetrical association have been proposed.[1,10] Theoretical analyses from thermodynamic considerations of oligomeric protein assemblies reveal that symmetrical associations are more energetically favorable compared to asymmetrical associations.[11,12] Additionally, an interface between subunits in a homooligomeric protein can offer new amino acid arrangements to facilitate the formation of a new active site.[13] Indeed, in evolution, forming an active site between subunits by reorganizing protein interfaces can be easier than generating a new active site within a single polypeptide by altering protein folding.[10] Further, oligomeric proteins also offer excellent opportunities for allosteric regulation.[14,15] Cooperative conformational changes between subunits in an oligomer allows efficient transfer of information from one subunit to the next or to all of the other subunits in a protein.[16]

1.1.2. Interfaces and Hotspots

Protein-protein interfaces govern many of the assembling and folding behaviors of oligomeric protein complexes.[17] When two or more protein monomers come into contact and several weak, noncovalent interactions occur among the amino acids at the surface of each monomer. Extensive studies of many protein-protein interfaces have revealed that these protein-protein interfaces recognize their binding partners with high specificity according to the complementarity of shape and chemical properties, such as forming hydrophobic patches, hydrogen bonds, or salt bridges.[16,18] Based on the shape and the chemical properties, interfaces of oligomeric proteins might be sorted into two categories: interfaces of obligate complexes and interfaces of nonobligate complexes.[19] However, some protein interfaces are hard to categorize clearly because oligomerization states often

depend on the conditions and environment.[20 21]

Three types of interactions (hydrophobic interaction, hydrogen bonding, and electrostatic interaction) are commonly found to be important in stabilizing protein-protein interfaces. Hydrophobic interactions often make large contributions to the stabilization of oligomeric protein interfaces.[22,23] When nonpolar amino acids interact with each other via van der Waals contacts in an aqueous environment, the exclusion of ordered water molecules from the contact area to the bulk solution provides a strong energetic driving force for association. Typically, protein-protein interfaces comprise a greater fraction of hydrophobic residues compared to noninterfacial protein surfaces. However, protein-protein interfaces also contain a fairly large amount of polar and hydrophilic residues and are therefore less hydrophobic than protein cores.[13] These polar groups form electrostatic and hydrogen bonding interactions that may also contribute to interface stability. Further, the specificity of protein association is largely determined by the complementary nature of electrostatic interactions and the geometric requirements of hydrogen bonding.[22] Indeed, on average, the number of hydrogen bonds is proportional to the interface area, occurring with a frequency of about one hydrogen bond per 200 $Å^2$ of protein interface surface area.[23,24] Additionally, interfaces can sometimes incorporate disulfide linkages or metal binding sites to gain extra stability.

Binding energies are not distributed evenly at protein interfaces. A group of a few residues that contribute a disproportionately large amount of the binding energy at intermolecular protein interfaces are called "hot spots."[25] Hot spots are usually clustered in interfaces that are highly complementary both in shape and amino acid arrangement. Although there is not a preference at hot spots for a single type of amino acid, such as

hydrophobic or charged residues, a systematic analysis of hot spots reveals a nonrandom amino acid composition; three amino acids (arginine, tryptophan, and tyrosine) are found to appear the most frequently at hot spots.[26] Tryptophan and tyrosine have bulky aromatic side chains, which can contribute to aromatic π-interactions, serve as hydrogen bonding partners and provide a large hydrophobic surface. Arginine is also capable of forming multiple types of favorable interactions, such as participating in up to five hydrogen bonds and forming a salt-bridge using its positively charged guanidino group. Isoleucine, histidine, and aspartic acid are also quite common in hot spots. Interestingly, compared to isoleucine, leucine is rarely found in hot spots even though they are constitutional isomers with identical chemical properties. Serine, threonine, and valine residues are also extremely rare at hot spots.[27]

Studying protein-protein interactions at the atomic level is crucial for understanding the biological functions and structures of proteins. It not only allows an extensive understanding of the chemistry of protein-protein interactions, it also can provide insight into protein self-assembly. The identification of important hot spot residues at protein-protein interfaces benefits many important applications, including rational drug design, analysis of metabolic and signal transduction pathways, and development of novel protein nanomaterials.

1.1.3. Hierarchical Assembly of Proteins

Numerous nanoscale architectures can be constructed from self-assembly of multiple protein building blocks.[2,28,29] For example, filamentous structures, such as amyloid-like fibrils and actin fibrils, can be obtained by assembling protein subunits linearly.[30] Such

linear assemblies can be made by two different types of subunit arrangements. In the first type, building blocks are assembled with the same direction relative to the filament axis and connected head-to-tail. In the second type, building blocks meet in a head-to-head and tail-to-tail assembly with two-fold rotational symmetry.

In addition to linear assemblies, polyhedral structures, such as viral capsids, can also be constructed by protein self-assembly.[31] Polyhedral capsids are assembled from multiple subunits in a highly symmetric fashion. Cubic symmetries have been proposed to be required for building an isomeric capsid that allows close packing of repeating subunits.[2] Three types of cubic symmetries, including tetrahedral (containing six two-fold axes and four three-fold axes), octahedral (containing 12 two-fold axes, eight three-fold axes, and six four-fold axes), and icosahedral (containing 30 two-fold axes, 20 three-fold axes, and 12 five-fold axes), are capable of generating capsids.

The hollow inner space of polyhedral capsids can be used to sequester guest molecules. Such encapsulation ability is used by viruses to protect their genomes. Some multi-enzyme complexes form cage-like structures that can be used to regulate enzyme activities by limiting the entry or exit of potential substrates.[32–34] In addition, by encapsulating a group of enzymes involved in a reaction cascade, the overall catalytic efficiency can be enhanced via substrate channeling.[35]

1.2. Polyhedral Capsids

1.2.1. Icosahedral Viral Capsids

Viruses construct protein shells that enclose their genomes to prevent damage or degradation. Many viral capsid shells are formed by arranging repeating subunits into

icosahedral symmetry. This strategy allows viruses to build a given size protein capsid shell economically by using only one or a few different types of subunits, which helps minimize genome size.[4]

Icosahedral symmetry requires 60 identical units with equivalent contacts between each of them. However, most icosahedral viral capsids are larger than a capsid that can be built by 60 identical subunits. To explain this observation, the rule of quasi equivalence and the lattice triangulation number T was introduced.[31] An icosahedron can be built from hexamers and pentamers constructed from more than 60 identical subunits, but are only quasi-symmetric, which means that the identical subunits are in different conformations and make different contacts with their neighbors in the icosahedron.

The facets of an icosahedron can been considered as 20 equilateral triangles, which can be further subdivided into smaller triangular subunits. This process is called triangulation.[36,37] According to the concept of the quasi equivalence rule, triangulation generates hexavalent (six-fold) and pentavalent (five-fold) positions on the icosahedral lattice. The triangulation process can be expressed by the equation

$$T = h^2 + hk + k^2 \tag{1.1}$$

In this equation, T is the triangulation number or the number of subunits in an icosahedral asymmetric unit. The vectors h and k correspond to the position of the five-fold vertex in the triangulated icosahedral lattice and they can be either zero or positive integers (Figure 1.1).

Quasi symmetry allows filling multiple identical subunits into each of the 60

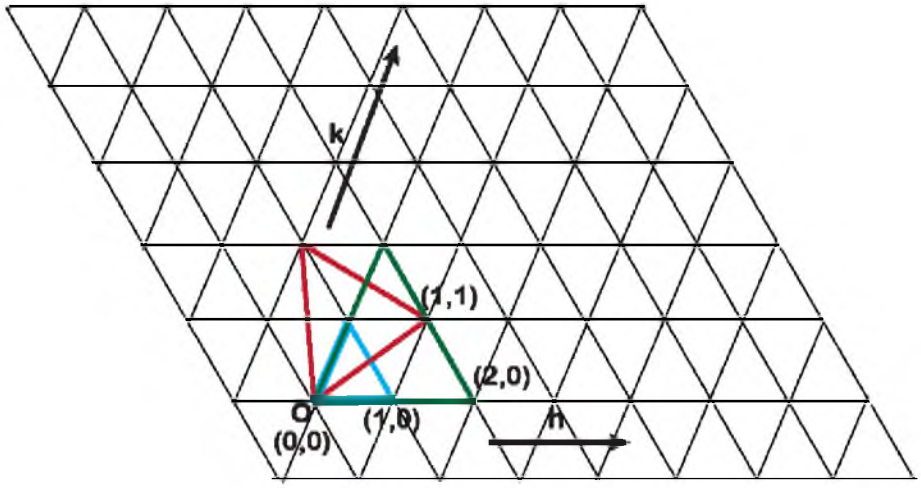


Figure 1.1. Triangulated facets for icosahedra. The geometry of each facet is described by the hexagonal coordinate system. The facets contain equilateral triangles with an integral number of subunits. The T number of an icosahedron is determined by the h and k vectors from the origin (O) and the equation, $T = h^2 + hk + k^2$. The facets of T = 1 (h = 1, k = 0), T = 3 (h = 1, k = 1), and T = 4 (h = 2, k = 0) icosahedron are shown as cyan, red, and green triangles, respectively.

symmetrically identical positions in the icosahedron (Figure 1.2). Consequently, the size of the 20 equilateral triangles and the icosahedral particle increase proportionally to the T number (under the assumption that the size of subunits stays constant). An icosahedral capsid with a given T number will consist of 60T subunits arranged into 12 pentamers with 10(T – 1) hexamers (Figure 1.3).

The spontaneous formation of capsid particles from multiple protein subunits is a sophisticated thermodynamic process driven by many weak protein-protein interactions.[38] Such an assembly process is similar to the classical nucleation theory of polymerization reactions for macromolecular polymers.[39,40] In this theory, assembly of a polymer is initiated by a nucleation event followed by an elongation step involving a cascade of faster reactions that add subunits into the polymer at stable rates until equilibrium between the polymer and the free subunits is reached. At the equilibrium stage, subunits can freely associate and disassociate from this open-ended polymer. A concentration (referred to as the critical concentration) of free subunits in solution at this stage must be present to maintain the polymer as a unique phase.

The overall assembly process of an icosahedral capsid exhibits sigmoidal kinetics.[38] There is a lag phase early in the assembly reaction, resulting from the slow formation of nuclei, followed by a rapid assembly of subunits and ending with a steady-state when the critical concentration of free subunits has been achieved.[41] The assembly pattern of an icosahedral capsid is similar to a classic polymer but with different kinetic and thermodynamic meanings. Icosahedral capsids assemble into a closed lattice with a defined number of subunits instead of an open-end polymer with an indefinite number of subunits. Further, in contrast to a classic polymer, in which the lag phase represents the

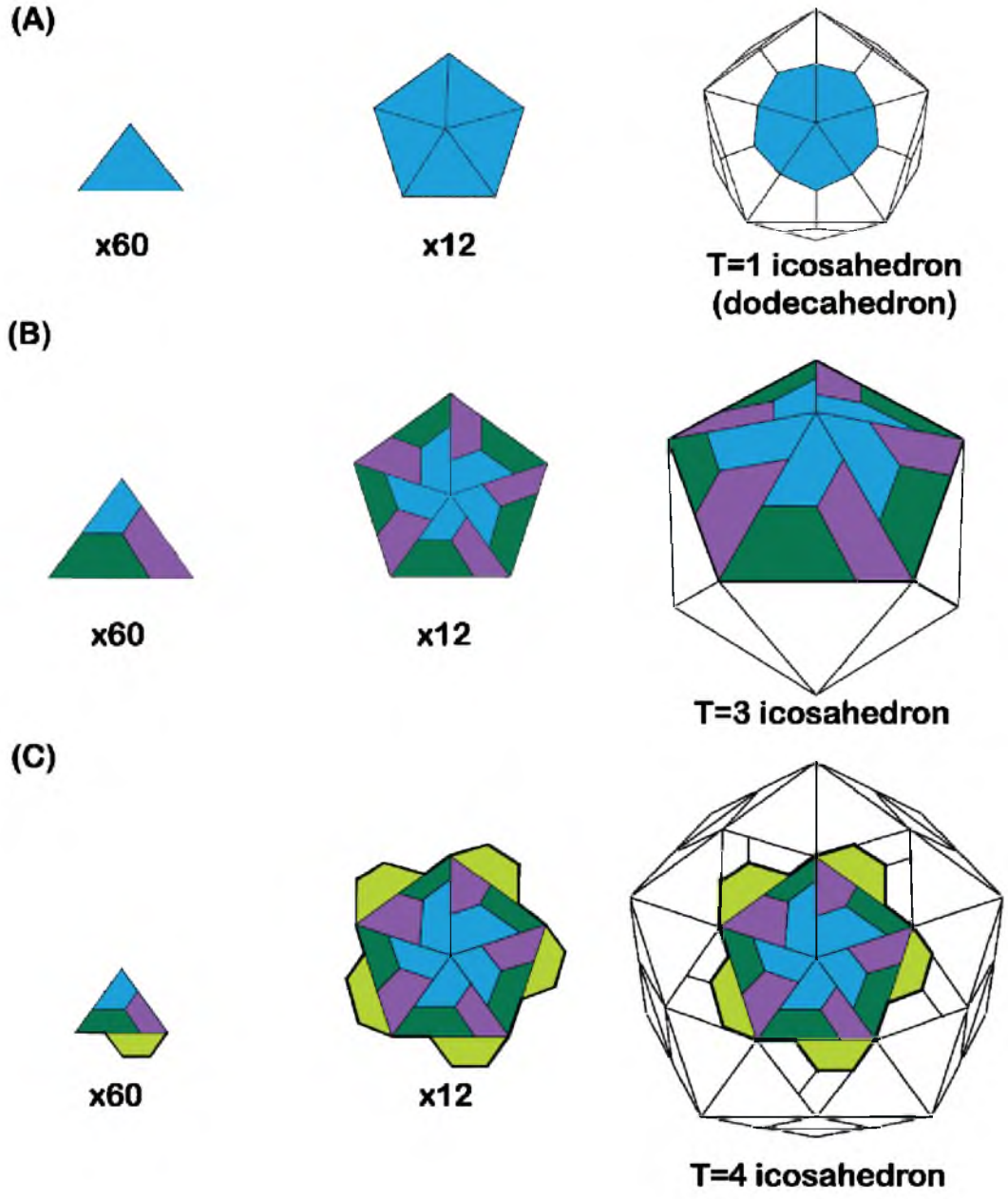


Figure 1.2. Models for T = 1, T = 3, and T = 4 icosahedral capsids. Quasi symmetry is shown for the T = 3 and T = 4 icosahedrons. Subunits in each asymmetric unit are colored differently. (A) The T = 1 icosahedral capsid consists of 12 pentameric building blocks for a total of 60 protein subunits. (B) The T = 3 icosahedral capsid consists of 12 pentameric and 20 hexameric building blocks for a total of 180 protein subunits. (C) The T = 4 icosahedral capsid consists of 12 pentameric and 30 hexameric building blocks for a total of 240 protein subunits.

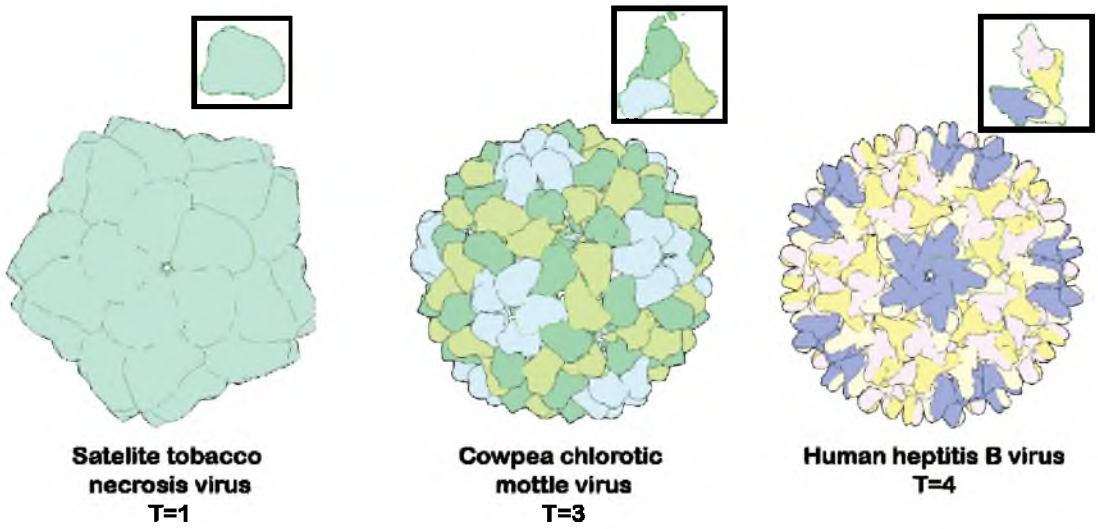


Figure 1.3. Examples of virus capsids with different T numbers. The asymmetric units of each capsid is shown in black squares; each subunit within an asymmetric unit is colored differently. The capsid structural models were generated using Chimera and the crystal structural information downloaded from RCSB protein data bank (www.pdb.org). The PDB id for satelite tobacco necrosis virus is 3s4g. The PDB id for cowpea chlorotic motte virus is 1cwp. The PDB id for human hepatitis B virus is 1qgt.

formation of a single nucleus, the lag phase of icosahedral capsid assembly corresponds to the generation of many individual nuclei. Each of these nuclei can go on to form sequentially larger and more complete intermediates. Complete capsids are then formed until the free subunit concentration reaches the pseudocritical concentration.[38]

1.2.2. Nonviral Protein Capsids

In addition to viral capsids, some nonviral proteins are also found to form polyhedral capsids in Nature. Many nonviral capsids form a reaction chamber by incorporating reactive molecules, such as enzymes, into their inner cavities. In some cases, proteins that form capsids can also function as catalysts, serving as an active participant in a multi-enzyme complex that carries out a reaction cascade. The compartmentalization of enzymes can regulate catalytic activity by controlling access of substrates to the enzymes inside of the capsid. Further, the catalytic efficiency can be enhanced by increasing the local concentration of enzymes and intermediates formed during multi-enzyme reaction cascades via substrate channeling. Examples of nonviral capsids that form multi-enzyme complexes include pyruvate dehydrogenase, fungal fatty acid synthase, lumazine synthase-riboflavin synthase complex, and bacterial microcompartments.[33,34,42–44]

Pyruvate dehydrogenase consists of three enzymes: pyruvate decarboxylase (E1 subunit), dihydrolipoyl acetyltransferase (E2 subunit), and dihydrolipoamide dehydrogenase (E3 subunit). The E2 subunits were found to self-assemble into 24-subunit octahedra (found in Gram-negative bacteria) or 60-subunit dodecahedral structures (found in eukaryotes and Gram-positive bacteria).[45] The E1 and E3 subunits noncovalently bind to the surface of the E2 capsid core. This E1/E2/E3 complex catalyzes

a directed sequence of acetyl coenzyme A biosynthesis reactions.

Fungal fatty acid synthase forms a 2600 kDa $\alpha_6\beta_6$ heterododecameric complex.[34] The cage-like structure of fatty acid synthase complex contains three distinct units. A central catalytic disk, formed by six α chains, is sandwiched by two domes, formed by the β chains, generating one reaction chamber on each side of the center disk. Six catalytic domains, which catalyze the sequential reactions of fatty acid biosynthesis, were found in the properly folded α and β polypeptide chains.

The multi-enzyme complex formed by *Bacillus subtilis* lumazine synthase encloses another enzyme, riboflavin synthase, in the interior of a T = 1 dodecahedroal capsid.[43] Lumazine synthase and riboflavin synthase together catalyze the last two steps in the riboflavin biosynthesis pathway. Thus, the encapsulation contributes to a more efficient catalysis at low overall substrate concentration.[46]

Bacterial microcompartments provide another example of multi-enzyme reactors that act as protein-bounded organelles.[44] The bacterial microcompartment capsid can enclose multiple types of enzymes, which catalyze a cascade of metabolic reactions, in the center of the capsids. There are several metabolic pathways carried out by different members of the bacterial microcompartment family.[47] Depending on the pathway, the members in the bacterial microcompartment family contain different sets of enzymes.[44] For example, the carboxysome, a polyhedral bacterial microcompartment that exists in most photosynthetic and chemoautotrophic bacteria, is involved in carbon fixation.[44] The carboxysome capsid simultaneously encapsulates two enzymes, ribulose bisphosphate carboxylase-oxygenase (RubisCO) and carbonic anhydrase (CA). CA catalyzes the dehydration of bicarbonate, which can be utilized by RubisCO that is also available in the

carboxysome chamber. The colocalization of CA and RubisCO results in increased local concentration of these enzymes and facilitates channeling CO_2 between them.[48] In addition, the carboxysome capsid can also function as a transportation gate to trap the CO_2 molecules inside of the capsid, providing a high local concentration of CO_2 by preventing escape via diffusion to the bulk environment.[47] As a result, the capsid forms an optimized environment for enhancing the efficiency of carbon fixation.

The capsids of bacterial microcompartments assemble from at least two different types of protein subunits into polyhedral structures with a size range of 80 to 200 nm.[35] The polyhedral capsids of bacterial microcompartments are mainly built by hexameric building blocks.[32] These hexagonal building blocks further self-assemble into tightly packed sheets, which form the flat facets of the polyhedral structure. A minor group of protein subunits found in the bacterial microcompartment capsids arrange into pentamers. Structural modeling and electron microscopy studies suggests that these pentamers form the vertices of the polyhedral structures.[49]

The tightly packed bacterial microcompartment capsid shell functions as an important diffusion barrier for controlling molecular transportation. Analysis of structural data led to the proposal that a narrow pore located at the center of each hexameric building block allows the transportation of small molecules into and out of the microcompartment.[32] Interestingly, an open form and a closed form have been observed for the pores in the hexameric building blocks.[50] These two dramatically different conformations perhaps provide a gated transportation mechanism of molecules across the protein shell. Other bacterial microcompartments such as those involved in propanediol utilization (Pdu) and ethanolamine utilization (Eut) also form polyhedral structures similar to the carboxysome

but they encapsulate different sets of enzymes involved in different metabolic pathways.[51,52]

A family of protein nanocompartments, encapsulin, forms ~20 nm particles in archaea and bacteria.[53] The encapsulin proteins purified from *Thermotoga maritima* self-assemble into T = 1 60-subunit icosahedral capsids. Encapsulin capsids specifically encapsulate ferritin-like proteins or peroxidases, both involved in oxidative-stress response. Interestingly, the guest enzymes possess a unique C-terminal extension motif, which drives selective encapsulation by binding to a highly conserved site on the interior surface of encapsulin. This C-terminal sequence perhaps can act as a tag to direct the encapsulation of guest proteins into the interior of encapsulin capsid.[53]

Ferritins, a family of protein capsids abundant in all domains of life, are well known to play an important role in iron storage and regulation.[54] Ferritins form octahedral capsids with 432 symmetry via self-assembly of 24 identical subunits.[55] While some enzymes which carry out important biological functions, such as electron transfer and oxygen transport, require iron for their activities, excess cellular iron can cause oxidative damage. Ferritins act as cellular iron detoxifiers and reservoirs by catalyzing the mineralization of iron into hydrous ferric oxide, which is stored in their inner cavities.[56,57] The pores at the three-fold axes in the protein shell of ferritins can function as gates to regulate the flow of iron across the shell.[58,59] The negatively charged amino acids at the three-fold pores create a negative electrostatic gradient that guides iron from the channel entrance to the active sites at the inner surface of ferritin capsids, where the Fe(II) becomes oxidized to Fe(III) by oxygen. The unique metal sequestration property of ferritins has been used to synthesize a variety of mineral cores within the protein shell.[60]

1.3. Engineering Protein Capsids for Biological and Material Applications

1.3.1. Nanoscale Material Synthesis

Some natural protein capsids, such as ferritins, catalyze biomineralization reactions and store inorganic particles in their inner cavities.[54] These protein capsids provide several advantages for nanoparticle synthesis. For example, the confined space within the protein capsids can be used for generating highly monodisperse nanoparticles. Also, these protein capsids provide a highly charged interior surface for inducing nucleation reactions to trigger mineralization.[59] A great deal of studies have utilized these protein capsids as nanosized bioreactors for the synthesis of nanoparticles within their cage-like structures.[60,61]

Engineered ferritin capsids have been used to synthesize a variety of inorganic and metallic nanoparticles including silver nanoparticles with a very narrow size distribution.[62,63] The inner surface of ferritin capsids was modified with dodecapeptides, which have been shown to reduce silver ions to metallic silver, allowing silver ions to be selectively reduced within the constrained space of the ferritin capsid interior and generating monodisperse silver nanoparticles. Other protein capsids, such as cowpea chlorotic mottle virus (CCMV) capsid, have also been engineered to produce inorganic nanoparticles.[64] Studies of ferritins reveal that the highly negatively charged interior surface of ferritin capsids can favor mineralization processes. Thus, the overall charge of the CCMV capsid interior surface was altered from positive to negative. The engineered CCMV capsids were demonstrated to generate iron oxide nanoparticles with homogeneous size distribution within the capsid.[65] This strategy could potentially be

applied to other protein capsids with different sizes to tailor-make nanoparticles with desired sizes.

The confined inner space of the heat shock protein (Hsp) capsid from *Methanococcus jannaschii* was used to synthesize defined polymer networks.[66] Cysteines were engineered onto the inner surface of the Hsp capsid to provide anchor points and initiate the polymerization reactions. Using an azide alkyne cycloaddition reaction approach, branched polymers were grown selectively within the Hsp capsid. This synthetic approach allows the synthesis of polymer structures in a highly restricted size range. In addition, the encapsulated polymer provided a high-capacity reservoir of small-molecule binding sites for drug delivery and bio-imaging applications.

1.3.2. Drug Delivery and Bio-imaging

Protein capsids are attractive scaffolds for developing novel drug carriers and have demonstrated promising results.[67] Their well-defined structures allow protein capsids to be modified by genetic or chemical methods without disrupting the overall quaternary structures. Further, functional molecules, such as antibodies, peptides, aptamers, carbohydrates, and polymers, can be conjugated onto the exterior surface of protein capsids to enable cell specific targeting, enhance cell uptake, or reduce immunogenicity.[68]

The hollow interiors of protein capsids allow capsids to encapsulate drugs or materials with optical or magnetic properties. In addition, the amino acids at the inner surface of protein capsids can be engineered to contain reactive sites, such as the reactive thiols on cysteines, for covalent bonding to cargo molecules. Alternatively, cargo molecules can also be loaded into capsids noncovalently via electrostatic interactions. By substituting

inner surface residues with positively or negatively charged residues, the loading efficiency of charged cargo molecules, such as DNA, RNA, or other ionic polymers, can be enhanced.

Numerous drug carrier systems have been generated using virus capsids as scaffolds. For example, the interior of bacteriophage MS2 capsid has been engineered to form covalent bonds with chemotherapeutic drugs, magnetic resonance imaging (MRI) contrast agents, positron emission tomography (PET) radioisotopes, and fluorescent dyes.[69–71] These therapeutic molecules were attached to the MS2 capsid either by forming a thioether linkage with the reactive sulfhydryl group on an engineered interior cysteine residue or by a diazonium coupling reaction with the interior tyrosines.

CCMV capsids also have been extensively studied as potential drug carriers. A metal-binding peptide was fused to the N-terminus of the CCMV capsid protein that allows the binding of gadolinium (Gd^{3+}), a MRI contrast agent.[72] Further, in another engineered CCMV capsid, a clinically relevant contrast agent, gadolinium-tetraazacyclododecane tetraacetic acid (GdDOTA), was covalently attached to lysine residues on the inner surface of the CCMV capsid protein via reaction with an N-hydroxysuccinimide (NHS) ester.[72]

Some virus capsids possess abilities to encapsulate negatively charged molecules, such as DNA, RNA, or synthetic polymers, in addition to their own genomic materials by forming capsid around the cargo molecules. Such abilities can be used to engineer drug carriers that encapsulate charged drug molecules noncovalently. For example, engineered hepatitis B virus capsids have been used to deliver synthetic DNA molecules to cell targets both *in vitro* and *in vivo*.[73] In addition, hibiscus chlorotic ringspot virus

capsid has demonstrated the ability to simultaneously encapsulate polyprotic acid-associated doxorubicin (an anticancer drug molecule) during capsid formation. In this study, folic acid was conjugated onto the drug-loaded capsids to promote selective cellular uptake of loaded capsids by cancer cells.[74] In another study, CCMV capsid was used to encapsulate DNA micelles that were self-assembled from DNA oligonucleotide amphiphiles.[75] Each DNA micelle possessed a hydrophobic core surrounded by anionic tails that point out from the core. The high negative charge density provided by the anionic tails of the DNA micelles allowed encapsulation into CCMV capsid. These DNA micelles can be preloaded with hydrophobic molecules that partition into the hydrophobic core or with hydrophilic molecules by sequence-specific hybridization to the tails of DNA micelles.

1.3.3.Catalysis

Protein capsids can be ideal scaffolds for developing novel bioreactors.[76] Pores in the capsid shell usually allow small molecules (such as substrates or products) to diffuse into and out of the capsid interior but keep big molecules (such as enzymes) inside the capsids. Indeed, bacterial microcompartments represent a successful example of natural bioreactors. Inspired by bacterial microcompartments, artificial nanobioreactors have been constructed using engineered protein capsids and have shown promising results as novel catalysts.[77–79] Further, the monodisperse inner space and the well-defined structure make protein capsids attractive as model systems for studying enzymatic reactions in a confined space.

The CCMV capsid has been engineered as a viral nanoreactor by encapsulating

horseradish peroxidase (HRP).[80] The assembly states of CCMV capsid can be dynamically switched between dimers and capsids at different pH values. For instance, CCMV protein forms a 180-subunit icosahedral capsid at pH 5 and the capsid dissociates into 90 dimers at pH values above 7.[81,82] Thus, HRP encapsulation can be achieved by mixing the disassembled CCMV proteins with HRP at pH 7 then subsequently bringing the pH down to 5 where the CCMV protein subunits assemble around the HRP molecules. Since there is no intrinsic affinity between HRP and the CCMV capsid, the encapsulation of HRP during CCMV capsid assembly depends on random diffusion, which results in a low yield of encapsulated enzymes. Consequently, those capsids that did take up HRP could be safely assumed to possess only one copy of the enzyme, which enabled single-molecule measurements of HRP kinetics. The observed activity revealed that the capsid shell is not only permeable to both substrates and products but can also regulate their diffusion across the capsid shell in response to pH-dependent changes in the size of the pores.[80]

In a follow-up study, a noncovalent anchoring approach was used to improve the control of guest encapsulation.[83] This strategy involved modification of both the CCMV capsid proteins and the guest enzymes with a positively charged peptide tag and a negatively charged peptide tag, respectively. Charge complementarity triggers the formation of a coiled-coil between the two peptide tags, driving encapsulation of the guest protein by the CCMV capsid. This strategy allows precise control over the number of encapsulated enzymes in each CCMV capsid.[77] The enzyme lipase B was encapsulated in this manner and was found to convert substrate to product more efficiently than free lipase B in bulk solution. However, increasing the number of enzyme molecules in the

capsid did not increase the overall reaction rate at a constant concentration of total enzyme.

In another example, cargo enzymes were packaged into bacteriophage Qβ particle via RNA-directed encapsulation *in vivo*.[79] In brief, a bifunctional RNA aptamer was engineered containing two binding domains, one to bind to the tagged cargo enzyme and one to bind to the capsid interior. Coproduction of capsid protein, guest enzyme, and RNA-aptamer led to the formation of capsids loaded with guest enzymes *in vivo*. The encapsulated enzymes were still catalytically active. Further, the protein capsid showed the ability to stabilize and protect the encapsulated enzymes against heat denaturation, protease degradation, and hydrophobic adsorption.

Bacteriophage P22 capsid was also used to generate a nanoreactor.[78] The P22 capsid is formed by self-assembling 420 copies of P22 capsid protein around a structural template formed by 100–330 copies of P22 scaffold protein, resulting in a T = 7 icosahedral capsid. By fusing the guest molecules with the scaffold proteins, guest can be encapsulated into P22 capsid during capsid assembly.[84] In this study, a truncated bacteriophage P22 scaffolding protein sequence was fused to the C-terminal end of the target enzyme, alcohol dehydrogenase D (AdhD). The P22 scaffolding protein-AdhD fusion protein promotes the specific encapsulation of cargo enzyme into the P22 capsid interior with a high loading yield during P22 capsid formation *in vivo*. Further, incubation of the P22 capsid at different temperatures can trigger the P22 capsid to transform into different capsid structures, including an expanded capsid structure and a "wiffleball" structure. Enzyme activity studies of the enzyme-loaded P22 capsids reveal that the overall activity of the encapsulated enzyme can be altered by transforming P22 capsid

into different morphological states, which change the internal volume and the porosity of the capsid.[78]

1.4. Switches for Capsid Assembly

1.4.1.Natural Capsid Assembly Switches

Protein capsids have been used to encapsulate different types of cargo molecules for many applications, as described above. For small cargoes, capsid loading can be carried out with intact capsids, but this strategy only works for cargoes that can fit through the capsid pores. However, the self-assembly/disassembly of some virus capsids can be controlled by changing solution conditions, allowing simultaneous cargo loading and capsid formation *in vitro*. Because such *in vitro* encapsulation approaches are usually performed in simpler and more controlled conditions, they could generate more homogeneous encapsulation products than *in vivo* encapsulation approaches. Further, *in vitro* cargo loading via controllable capsid assembly could expand the range of potential cargo molecules, increasing the versatility of capsids for bionanotechnological applications.

The pH-dependent assembly of CCMV capsid, which has been discussed above, provides a well-known example of natural capsid assembly switches. The assembly states of CCMV capsid can be freely exchanged between dimers and 180-subunit icosahedral capsids by changing the buffer pH.[81] In another example, hepatitis B virus capsid also exhibits an assembly switch, which can be induced to assemble by increasing the ionic strength. Wild-type hepatitis B virus capsid consists of 240 identical protein subunits, and these subunits are arranged into a T = 4 icosahedral capsid. A truncated hepatitis B virus

capsid protein forms dimers after purification from *E. coli* cells. These dimers reassemble into wild-type-like capsids in response to high ionic strength. Interestingly, mild denaturing conditions (1.75 to 3.25 M of urea) can be used to trigger the dissociation of hepatitis B virus capsids without unfolding the dimeric building blocks.[85,86]

Capsid assembly can also be induced upon metal ion binding. Studies on icosahedral murine polyomavirus (MPV) capsid demonstrated that MPV capsid self-assembly from pentameric building blocks can be triggered *in vitro* by adding Ca^{2+} ions.[87] Further, lowering the pH or increasing the ionic strength can also induce the self-assembly of MPV capsids. Although a detailed understanding of these natural capsid assembly switches is lacking, protein-protein interactions at the interfaces of building blocks should be crucial in determining the assembly states of protein capsids. Thus, knowledge of protein-protein interactions could help further our understanding of virus capsid assembly and aid in the development of artificial capsid assembly switches.

1.4.2. Engineering Capsids for Controlled Assembly

The use of natural capsid assembly switches can be a very attractive strategy for cargo loading. However, not all protein capsids possess this feature. Further, the solution conditions that trigger natural assembly switches may sometimes be incompatible with the desired application. Such limitations could be overcome by using rationally designed artificial capsid assembly switches as alternatives. Previously reported attempts to design artificial capsid assembly switches exploit various triggers, including metal ion-induced capsid assembly, pH-dependent capsid assembly, or temperature-dependent capsid assembly.[88–91]

1.4.2.1. Metal-induced Capsid Assembly

Some protein oligomerizations can be controlled upon binding of small molecules or metal ions across the protein-protein interfaces. Such strategies could be useful for designing a potential switch to regulate assembly of a protein capsid. Ferritin capsid, which is lacking a natural capsid assembly switch, was used as a model system for engineering artificial capsid assembly switches.[88] Metal-induced ferritin capsid formation was achieved by engineering two metal coordination sites, which consisted of two pairs of histidine residues for each binding site, at the two-fold axes of the ferritin capsid. Each binding site is capable of forming a planar coordination with one copper ion. The monomeric building blocks used for studying the metal-induced capsid formation also included three additional mutations of the interfacial residues in order to disrupt native hydrogen-bonding interactions, which stabilize the wild-type ferritin capsid formation. The resulting ferritin variant, called MIC1, was purified from inclusion bodies as monomers in the presence of EDTA. Upon removal of EDTA and addition of copper ions, MIC1 was converted into a 24-subunit capsid. However, the metal-induced MIC1 capsid did not disassemble upon removing the metal ions with EDTA.

In another example, CCMV capsid was also engineered to allow metal-induced capsid assembly.[91] As described previously, CCMV capsid proteins can be reversibly assembled and disassembled by changing buffer pH. The usefulness of this switch is limited by the fact that the capsid disassembles into protein dimers at physiological pH. Therefore, an assembly switch that can stabilize the CCMV capsid at high pH could be valuable. In the metal-induced CCMV capsid assembly study, the N-termini of the CCMV subunits, which are located at the six-fold and five-fold axes of the capsid, were fused to

hexahistidine sequences, which provided a metal coordination motif. Polyhistidine sequences form strong multivalent interactions with metal ions; therefore, this modification is commonly used for protein purification by affinity chromatography. The N-terminal hexahistidine modified CCMV proteins, named His-CPs, formed capsid-like assemblies at pH 7.5 when nickel ions were added to the buffer. The nickel induced CCMV assemblies can be disassembled by adding EDTA to remove the nickel ions.

1.4.2.2. pH-dependent Capsid Assembly

Some protein capsids, such as CCMV capsid, have an intrinsic ability to switch their assembly states in response to changing the environmental pH. Inspired by this type of naturally occurring switch, the protein interfaces of the E2 subunit in pyruvate dehydrogenase capsid were modified to enable pH-responsive disassembly.[90] The pyruvate dehydrogenase capsid is formed by 60 E2 subunits, which self-assemble into a dodecahedral capsid with 532 symmetry. A cluster of six histidine residues was found at the three-fold axis of wild-type E2 capsid. Since the imidazole group of histidine typically has a pK_a value of ~6–7, histidine is mainly neutral at physiological pH but can be protonated and carry positive charge when the pH is lowered below 6.5.[92,93] Therefore, the histidine residues at the three-fold axis of the E2 capsid can be used to introduce charge-charge repulsion at low pH condition. The wild-type E2 forms stable capsids at both pH 7.4 and pH 5.0. However, after deleting the first 44 amino acids from the N-terminus, the truncated E2 protein (ΔN-E2) assembled into capsids upon production in *E. coli* and isolation at pH 7.4, but the capsids disassembled at pH 5.0, forming insoluble aggregates.

1.4.2.3. Temperature-dependent Capsid Assembly

CCMV capsid has also been engineered for temperature-dependent assembly by attaching temperature-responsive elastin-like polypeptides (ELPs) to the capsid protein.[91] ELPs contain a repeating Val-Gly-Xaa-Gly sequence, in which the Xaa can be any natural amino acid except proline. ELPs have shown reversible switching from an extended water-soluble state to a hydrophobic state in response to the temperature.[94,95] The transition temperature of ELPs can be varied by multiple factors such as the residue at the Xaa position, the number of repeats, and the protein and salt concentrations. In the temperature-dependent CCMV capsid assembly study, an ELP sequence was fused to the N-terminus of the CCMV capsid protein. The resulting protein, ELP-CP, was able to form a wild-type-like T = 3 capsid at pH 5.0, but gave only a small yield of capsids in pH 7.5 buffer containing 1.3 M NaCl at 20 °C. However, upon incubation at 35 °C for 15 min, most of the ELP-CP proteins assembled into T = 1 capsids.

1.5. Lumazine Synthase

Lumazine synthase can be an interesting model system for studying protein capsid self-assembly because both capsid-forming and noncapsid forming lumazine synthase homologs have been observed. Lumazine synthase is found in bacteria, archaea, fungi, and plants, and catalyzes the formation of 6,7-dimethyl-8-ribityllumazine, the immediate precursor of riboflavin, from 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione and 3,4-dihydroxy-2-butanone 4-phosphate.[96] Riboflavin synthase catalyzes the addition of two 6,7-dimethyl-8-ribityllumazine molecules to form riboflavin and 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione (Figure 1.4). A noteworthy feature of this two-

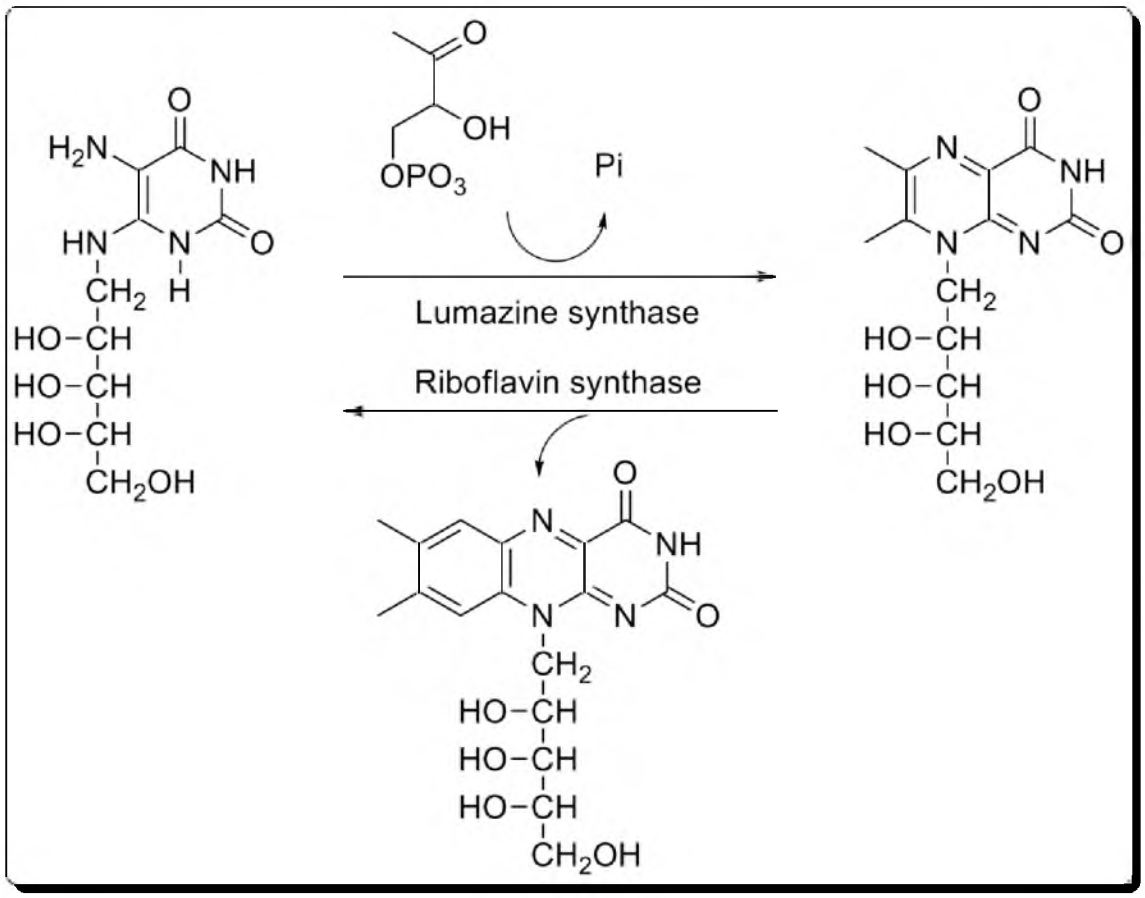


Figure 1.4. The reactions catalyzed by *Bacillus subtilis* lumazine synthase/riboflavin synthase complex.

step process is the cycling of material between the two enzyme active sites.

Lumazine synthase possesses various quaternary structures among different organisms, although the tertiary fold is highly conserved. In some fungi and bacteria, such as *Saccharomyces cerevisiae*, *Magnaporthe grisea*, *Candida albicans*, *Schizosaccharomyces pombe*, and *Mycrobacterium tuberculosis*, lumazine synthase forms a pentameric arrangement of five identical monomers (Figure 1.5A).[97–101] Lumazine synthase from *Brucella abortus* assembles as a D_5-symmetric dimer of pentamers (Figure 1.5B).[102] Interestingly, lumazine synthases from some other bacteria and plants, such as *Bacillus subtilis*, *Escherichia coli*, *Aquifex aeolicus*, *Salmonella typhimurium*, and *Spinacia oleracea*, form 60-subunit capsids (Figure 1.5C).[99,103–106] Typically, lumazine synthase capsids consist of 12 pentameric building blocks, arranged into T = 1 dodecahedra with outer diameters of about 160 Å. However, multiple assembly states of *B. subtilis* lumazine synthase (BsLS) capsid have been observed in different buffer systems, at different pH values, and with different ligands. Besides T = 1 capsids, expanded T = 3 capsids, with a diameter around 29 nm, and even T = 4 capsids of BsLS have been observed by using transmission electron microscopy (TEM).[107]

BsLS forms a multi-enzyme nanoreactor by encapsulating a riboflavin synthase trimer in the inner space to enhance catalytic efficiency via substrate channeling.[43,46] An approximately 16 Å pore (C_α to C_α) located at the five-fold symmetry axis of the lumazine synthase pentamer has been proposed as the route by which substrates and products diffuse across the capsid shell and access the active sites at the subunit interfaces within the pentamers.[99] However, structural analysis indicates that this pore is too narrow to permit diffusion of the product, riboflavin, out of the capsid. Interestingly, an *A. aeolicus*

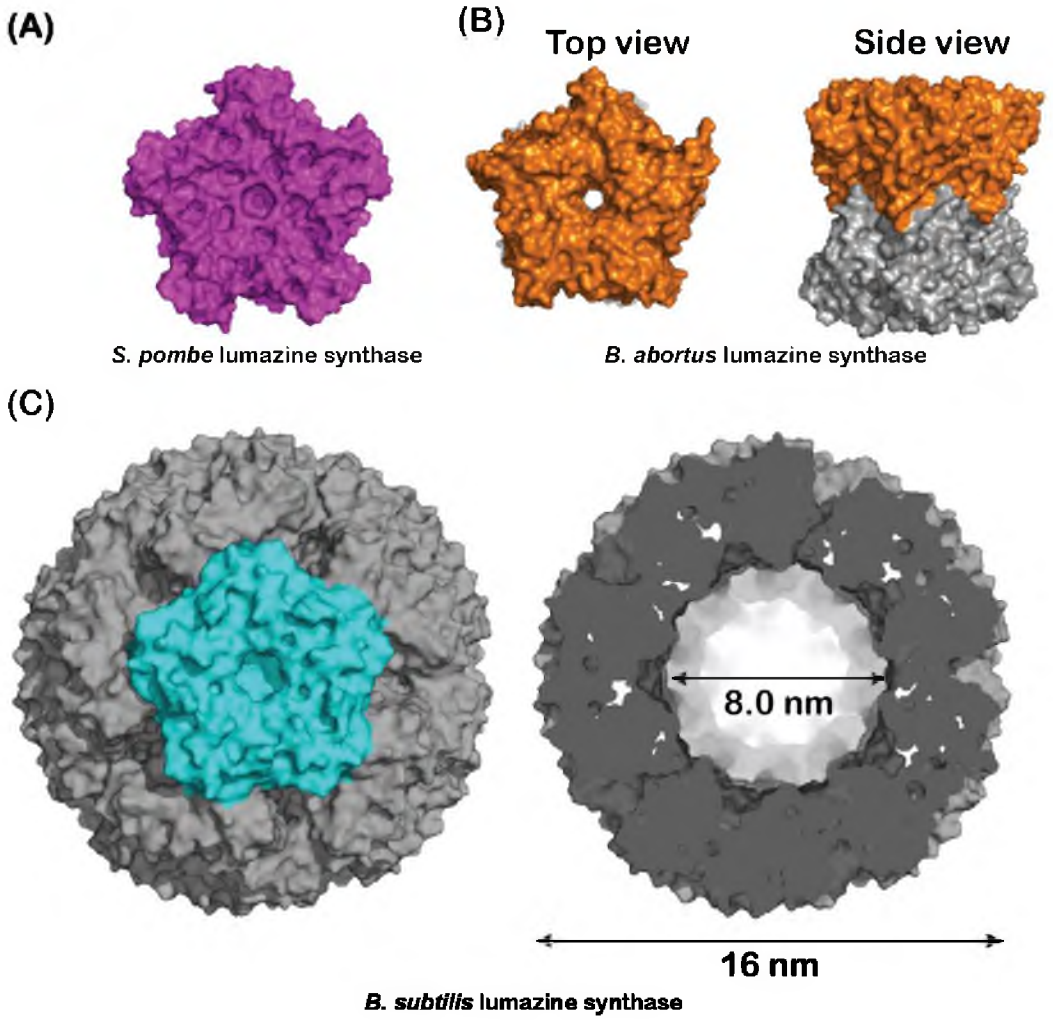


Figure 1.5. Structural comparison of lumazine synthase family members. (A) Surface representation of *S. pombe* lumazine synthase pentamer (PDB ID:1KYV) viewed along the five-fold symmetry axis. (B) Surface representation of *B. abortus* lumazine synthase dimer-of-pentamers (PDB ID: 1XN1). One pentamer was colored orange and the other was colored grey. (C) left: surface representation of *B. subtilis* lumazine synthase capsid (PDB ID: 1RVV) viewed down the five-fold symmetry axis. One of the 12 pentameic building blocks was colored as cyan, the others were colored grey. Right: cross-section of the *B. subtilis* lumazine synthase capsid, showing the inner and outer diameters

lumazine synthase (AaLS) variant, which contains a four amino acid insertion, homologous to a sequence found in pentameric lumazine synthase, forms an expanded T = 3 capsid.[107] This altered assembly demonstrates the plasticity of protein-protein interfaces in the capsid. Cryo-electron microscopy studies suggest that this expanded capsid can be a consequence of self-assembling expanded pentameric building blocks with widened pores in the centers of the pentamers.[108] This expanded pentameric structure provides a possible model for an alternative conformation of the lumazine synthase pentamer. Based on this observation, a model of capsid breathing during the catalytic cycle of lumazine synthase was proposed.[108] This model proposed that the expanded T = 3 capsid is formed by expanded pentamers with unoccupied substrate binding sites. The expanded pentamers in the T = 3 capsid contain widened pores that allow transport of substrates across the capsid wall to the active sites. Upon binding of substrates, the pentamers change to a closed state and form a T = 1 capsid. After the catalytic reaction is done, the products are released from the active site and the pentamers revert back to the expanded state with widened pores, which allows the products to diffuse out of the capsid interior.

Lumazine synthase capsids have been engineered as molecular containers for a variety of molecules.[109,110] For example, a BsLS based bionanoreactor has been utilized to synthesize iron oxide nanoparticles.[110] One homolog of BsLS, AaLS, is a particularly promising scaffold for engineering novel molecular containers.

The AaLS capsid possesses many desirable features for protein engineering. One important property of AaLS is its extremely high thermostability, with a melting temperature of 119.9 °C, which likely stems from the fact that this protein is derived from

a hyperthermophilic bacterium.[103] Further, the AaLS capsid can be produced recombinantly in *E. coli* with high yields. An additional advantage for protein redesign is the availability of a high-resolution crystal structure. Indeed, AaLS has been engineered as a molecular container for both protein and RNA guests, based on charge complementarity.[111–114]

To encapsulate proteins, four residues per monomer (R83, T86, T120, and Q123), located at the inner surface of the AaLS capsid, were simultaneously mutated to glutamates to generate a variant named AaLS-neg.[111] Surprisingly, AaLS-neg forms a 180-subunit T = 3 capsid that has 720 additional negative charges relative to wild-type AaLS. Charge complementarity between AaLS-neg and a cationic R_{10} tag fused to the green fluorescent protein (GFP) provided affinity between the capsid and this model guest protein. Coproduction in *E. coli* resulted in the encapsulation of three to four GFP-R_{10} molecules per AaLS-neg capsid.

The loading capacity of R_{10}-tagged guest proteins in the AaLS-neg capsid was further optimized by a directed evolution approach. In this experiment, the R_{10}-tag was fused to the C-terminus of the cytotoxic HIV protease. Coproduction of the tagged protease and AaLS-neg conferred a growth advantage on the cells relative to a control coproducing the tagged protease and wild-type AaLS. Presumably, confinement in AaLS-neg capsids prevents HIV protease from destroying cellular proteins, which enabled the selection of capsid variants that more efficiently encapsulate this toxic enzyme. After four rounds of selection, an AaLS variant, AaLS-13, was identified that showed an almost 10-fold higher encapsulation efficiency than the parent AaLS-neg.[112] The increased encapsulation efficiency likely stems from a further increase in the number of negatively charged

residues on the inner capsid surface.

AaLS-13 has also shown the ability to encapsulate supercharged GFP(+36), a variant of GFP in which 29 surface-exposed residues were mutated to positively charged amino acids, *in vitro*. By optimizing the concentrations of capsid and guests, up to 100 GFP(+36) molecules can be loaded per capsid. When produced in *E. coli* cells without a cationic guest, AaLS-13 purifies as a mixture of pentamers, higher order oligomers, and fully assembled capsids. Interestingly, the encapsulation of GFP(+36) *in vitro* works equally well with either intact AaLS-13 capsids or lower-order assemblies. The former result suggests that the assembly state of empty AaLS-13 capsids may be dynamic. The latter result suggests that cationic guests can act as templates for AaLS-13 capsid assembly.

In contrast to AaLS-neg and AaLS-13, another AaLS variant, AaLS-pos, was engineered with a positively charged inner surface. In this case, charge complementary was used to encapsulate cellular RNAs.[113] AaLS-pos contains four amino acid substitutions per monomer (T86R, D90N, T120R, and E122R) yielding a net interior charge of up to +240, providing a much more highly cationic environment compared to wild-type capsid, which is expected to possess a net interior charge of -60. After production in and purification from *E. coli* cells, AaLS-pos forms 60-subunit T = 1 capsids containing cellular RNA molecules ranged in size approximately from 50 to 400 bases in length. AaLS-pos represents a promising starting point for the development of novel nucleic acid encapsulation systems which could be used as delivery vehicles for therapeutic molecules, such as siRNAs.

1.6. Aims of This Thesis

The work presented in this thesis seeks a more detailed understanding of the interactions involved in the self-assembly of AaLS capsids in order to extend its functional capabilities. The results should contribute to the long term goals of this project, namely the design of novel AaLS encapsulation systems useful for developing drug delivery systems and enzymatic nanoreactors.

Here, the AaLS capsid is viewed as a dodecamer of pentamers, in which the interactions across the pentamer-pentamer interfaces are essential for capsid formation but not the stability of individual pentamers. Chapter 2 describes a detailed analysis of important interaction sites at the AaLS pentamer interfaces. Consequently, AaLS assembly can be halted at the pentamer stage by making a small number of well-chosen pinpoint mutations that alter interactions that stabilize the interfaces between pentamers.

Using knowledge of important interactions at the pentamer interfaces in the AaLS capsid, a redox-based assembly switch was designed, constructed, and characterized, as outlined in Chapter 3. An engineered pentameric AaLS variant, AaLS-switch-red, in which a unique cysteine was introduced at a position corresponding to the three-fold symmetry axis of the capsid, was used as a starting point for capsid assembly *in vitro*. Upon modification with thiophenol, the resulting adduct, AaLS-switch-ox, was found to self-assemble into 60-subunit capsids, presumably by mimicking a hydrophobic interaction present in wild-type capsids.

To better control AaLS capsid self-assembly *in vitro*, a pH-dependent capsid assembly switch that can reversibly interconvert between assembly states was designed, constructed, and characterized, as described in Chapter 4. Three histidine mutations per

polypeptide were introduced near the three-fold symmetric pentamer interface. The titratable histidine side chains were used to generate positive charge at low pH, driving capsid disassembly by causing charge repulsions that destabilize the pentamer interface. The capsids could be reassembled by increasing the pH back to its original value since these histidines minimally affect capsid assembly when neutral.

The results in this thesis demonstrate the ability to control AaLS capsid assembly *in vitro*, which could expand the range of potential cargo molecules for this scaffold. These controllable assembly switches enhance the versatility of the AaLS capsid and pave the way for its use in many applications, including the development of bio-imaging agents, nanocontainers, or drug delivery systems.

CHAPTER 2

CONVERSION OF A DODECAHEDRAL PROTEIN CAPSID INTO PENTAMERS VIA MINIMAL POINT MUTATIONS

2.1. Introduction

Many proteins form symmetrical, nanoscale architectures from multiple copies of identical polypeptide subunits.[1] These assemblies are held together by noncovalent interactions at the protein-protein interfaces.[13] However, the energetic contributions of individual residues to the stability of protein-protein interfaces are often unevenly distributed.[25,27] Subtle changes can sometimes convert one quaternary structure into another. General rules for predicting quaternary assembly have been slow to emerge.[115–117]

The engineering of nanostructures demands a thorough understanding of the basis for protein self-assembly.[29] In particular, self-assembled protein capsids represent a class of nanoscale scaffold that holds much promise for various applications, including drug delivery, catalysis, and materials synthesis.[67,76,118,119] Indeed, much effort has gone into repurposing capsids to act as containers for a wide variety of nonphysiological cargo molecules, such as proteins, nucleic acids, small molecules, and inorganic nanoparticles.[75,79,83,84,120–125] The controlled and efficient loading of these containers can be facilitated by the ability to interconvert the capsid with lower order quaternary states.[83,126] Viral capsids possess inherent switches that allow the efficient disassembly and reassembly in response to solution conditions.[127,128] In principle, it should also be possible to engineer the interconversion of quaternary structures between capsids and lower order assemblies.

The capsid formed by lumazine synthase from *Aquifex aeolicus* (AaLS) is a promising scaffold for engineering novel molecular encapsulation systems.[111,112,114] AaLS self-assembles into a 60-subunit dodecahedral capsid (Figure 2.1A) with a diameter

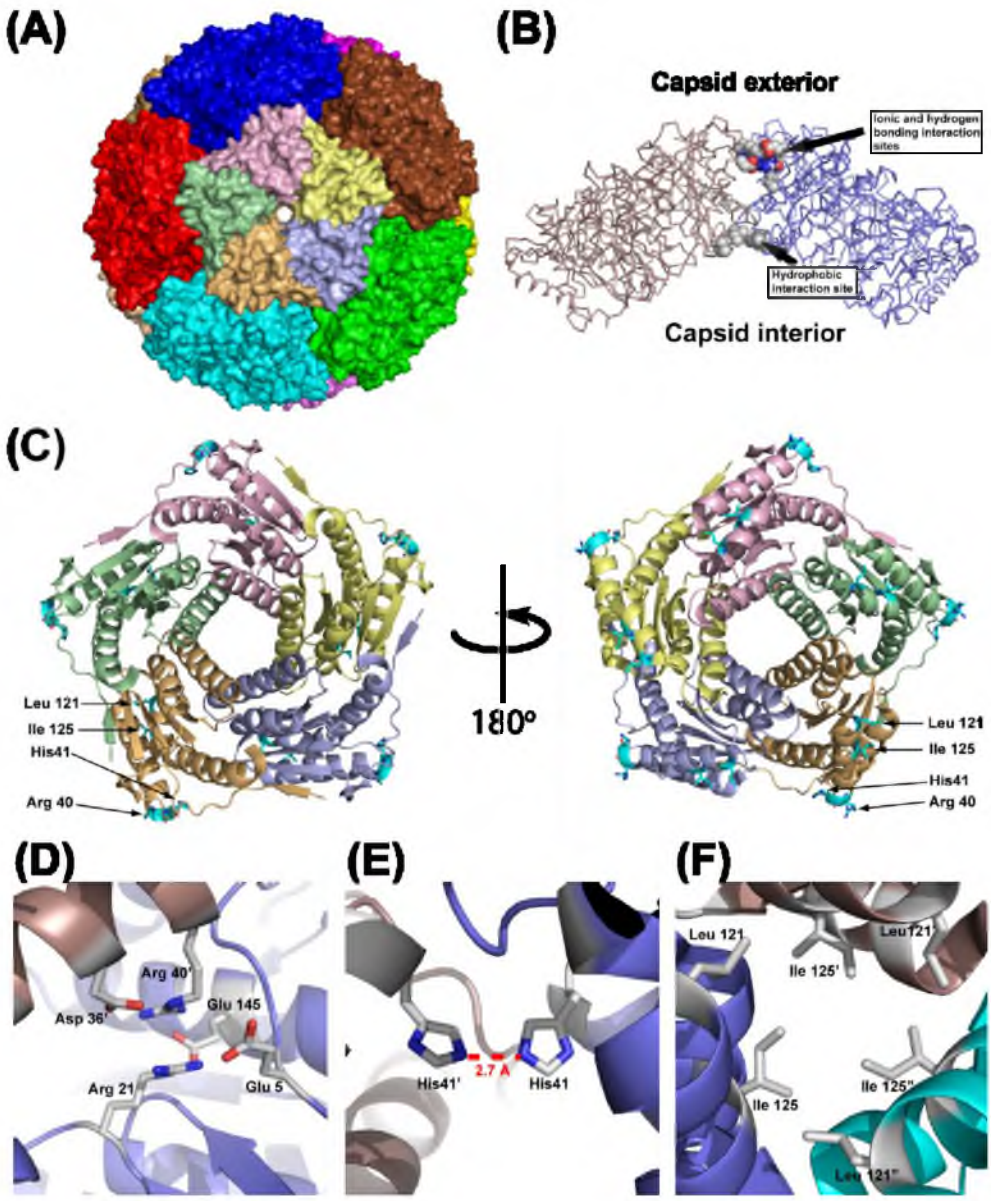


Figure 2.1. Structure and protein-protein interactions of the AaLS capsid. (A) Surface representation of the AaLS capsid. All of the pentamers are colored differently. In the pentamer at the front of the capsid only, each subunit is colored individually. The rest of the pentamers are colored uniformly. (B) Side-view of two adjacent pentamers extracted from the crystal structure of the AaLS capsid. The backbones of the pentamers are shown in brown and blue. Residues comprising the ionic, hydrogen bonding, and hydrophobic interaction sites are shown in space-filling representation. (C) Ribbon diagram of a pentamer extracted from the crystal structure of AaLS capsid. The view on the left is from the outside of the capsid, and the view on the right is from the inside of the capsid. Amino acids that contribute to the ionic, hydrogen bonding, and hydrophobic interaction sites are colored cyan and indicated by arrows. (D) The ionic interaction site. (E) The hydrogen bonding site. (F) The hydrophobic interaction site.

of 16 nm, enclosing an empty chamber approximately 9 nm across.[103] This capsid is extremely stable, with a reported melting temperature of 120 °C. While the AaLS capsid presumably assembles from 12 pentamers, lone pentamers of AaLS have never been isolated.

Despite an inability to reversibly disassemble AaLS *in vitro*, this capsid has been engineered to encapsulate guest proteins *in vivo* using a charge complementarity strategy.[111,112] In this system, AaLS variants with extra negative charges at the inner surface of the capsid act as hosts for guest proteins bearing a positively charged decaarginine (R_{10}) tag. Encapsulation complexes are formed by coproduction of the AaLS variant and guest protein in *Escherichia coli* cells. Loaded capsids can be purified from these cells, but are essentially dead-end complexes. For example, a rationally designed AaLS variant (AaLS-neg) in which four residues at the inner capsid surface have been changed to glutamates can encapsulate R_{10}-tagged green fluorescent protein or HIV protease .[112] Interestingly, the addition of negatively charged residues at the inner surface of the capsid causes a dramatic expansion of the capsid structure, as AaLS-neg forms 180-subunit capsids with a diameter of 28 nm. Directed evolution of AaLS-neg to optimize the encapsulation of R_{10}-tagged HIV protease gave a variant (AaLS-13) with increased negative charge density at the inner surface that formed capsids 34 nm in diameter, over twice the size of wild-type AaLS.[112] When produced in the absence of the HIV protease guest, AaLS-13 is isolated a mixture of intact capsids and capsid fragments. These fragments can assemble into capsids in the presence of positively charged guest protein.[114] These observations suggest that it should be feasible to tip the balance between different assembly states of AaLS in a controlled fashion.

In Nature, the lumazine synthase (LS) family of proteins is rich in quaternary structural diversity. In addition to 60-subunit capsids, many homologs form pentamers.[97,99] Decamers have also been observed that assemble via the dimerization of pentamers.[129] Various explanations for what determines the extent of assembly, in particular capsid versus pentamer, have been put forward based on homology. For example, phylogenetic and structural analysis led to the identification of eight sequence positions that have different patterns of conservation between capsid-forming and pentameric LSs.[130] Structural comparisons have also highlighted potential suspects responsible for determining the extent of assembly. Persson *et al.*[99] purported that a proline residue near the N-terminus of pentameric LSs could prevent capsid formation by breaking a β-strand that aligns with a β-sheet of an adjacent subunit in capsid-forming LSs. Alternatively, Braden *et al.*[131] suggested that the presence of a five-residue kink in the C-terminal helix is important for capsid formation. In addition, it was also proposed that a four-residue insertion in the loop connecting the fourth and fifth α-helices of a pentameric LS could be responsible for preventing further assembly.[97] This latter hypothesis was tested experimentally by making the insertion in AaLS.[107] The resulting variant still formed capsids, albeit with an expanded structure.

Here, we aimed to identify structural determinants of capsid formation in AaLS by targeted mutation of residues that define the pentamer-pentamer interfaces of the capsid. Using visual inspection of the available high-resolution crystal structure, we identified specific noncovalent interactions between side chains on adjacent pentamers in the AaLS capsid. These interactions were disrupted by strategic amino acid substitutions, effectively preventing capsid formation by halting assembly at the pentamer stage.

2.2.Results

2.2.1. Design of AaLS Variants

The AaLS capsid can be viewed as a dodecamer-of-pentamers (Figure 2.1A). The hierarchical assembly of the capsid from pentameric building blocks is further supported by the observation that the assembly of some homologs halts at the pentamer stage.[97,99] To understand the determinants of capsid formation in AaLS, we examined the interface between pentamers in the available crystal structure (PDB ID: 1hqk).[103] Three distinct types of interaction sites (hydrogen bonding, ionic network, and hydrophobic cluster) are found that presumably hold the pentamers together in the capsid (Figures 2.1B–F).

The hydrogen bonding interaction found at the two-fold symmetry axis of the capsid is formed by two histidine residues, H41 and H41', from two adjacent pentamers (Figure 2.1E). The length of this hydrogen bond is 2.7 Å, which marks the point of closest approach between two pentamers. The ionic network, located next to the two-fold symmetry axis, is highly conserved in dodecahedral lumazine synthases and has been previously reported to play an important role in AaLS capsid stability.[107] Two arginine residues, R21 and R40', from two different pentamers lie at the center of this network (Figure 2.1D). Their guanidino side chains are stacked on top of each other at a distance of 3.3 Å and are surrounded by the acidic residues E5, E145, and D36'. The third pentamer interaction site involves residues L121 and I125, which form a hydrophobic cluster at the three-fold symmetry axis interface near the inner surface of the capsids (Figure 2.1F).

We hypothesized that we could change the AaLS assembly state from a 60-subunit capsid to a pentamer by mutating a subset of the residues listed above. To test this hypothesis, a series of AaLS variants were designed to achieve the targeted disruption of

the noncovalent interactions at the interface between pentamers in the wild type AaLS capsid. Among the amino acids listed above, we focused on five amino acids (R21, R40, H41, L121, and I125) in particular. Substituting negatively charged glutamates in place of these amino acids should both abolish stabilizing interactions and introduce charge-charge repulsion at the native pentamer-pentamer interface. Alternatively, introducing arginine residues at some of these positions could achieve the same effect while also potentially generating steric clashes.

2.2.2. Screening Assembly State by SEC

A series of AaLS variants containing point mutations designed to prevent the formation of stabilizing noncovalent interactions between pentamers were produced in BL21 DE3 *E. coli* cells and purified by ammonium sulfate precipitation followed by anion-exchange chromatography. The assembly states of these AaLS variants were estimated by SEC with a sephacryl S-400HR column (Tables 2.1 and 2.2). As suggested by our AaLS capsid structure analysis, R40, H41, L121, and I125 all contribute to capsid assembly. Indeed, simultaneous mutation of the ionic, hydrogen bonding, and hydrophobic interaction sites (R40E/H41E/L121E and R40E/H41E/I125E) appeared to give stable pentamers (Figure 2.2).

Mutating just the ionic and hydrogen bonding sites (R40E/H41E and R21E/R40E/H41E) has no effect on the capsid assembly. Interestingly, changing both residues at the hydrophobic site (L121E/I125E) causes the capsid to expand, which may reflect partial destabilization. This expanded capsid species could contain 180-subunits, as has been seen in other variants of AaLS.[107,111] On the other hand, simultaneous

Table 2.1. Assembly states of AaLS variants.

Protein	# of pentamers[a]
wild type AaLS	12
wild type ScLS	1
R40E/H41E-AaLS	12
R21E/R40E/H41E-AaLS	12
L121E/I125E-AaLS	36[b]
R40E/L121E/I125E-AaLS	2 or 3[c]
H41E/L121E/I125E-AaLS	2 or 3[c]
H41R/L121R-AaLS	1
R40E/H41E/L121E-AaLS	1
R40E/H41E/I125E-AaLS	1
R40S/H41S/I125S-AaLS	1
R40S/H41L/L121E-AaLS	1

[a] Determined by SEC using a sephacryl S-400HR column. The number of pentamers is defined as the total number of protein subunits divided by five.

[b] This protein eluted similarly to AaLS-neg, which forms a 180-subunit capsid. Thus, this variant probably assembles from 12 pentamers and 20 hexamers, rather than from 36 pentamers.

[c] These proteins eluted at volumes between those expected for a dimer-of-pentamers and a trimer-of-pentamers.

Table 2.2. SEC peak elution volumes for the AaLS variants.

Protein	S-400 elution volume (mL)	S-300 elution volume (mL)
Wild type AaLS	69.4	38.9[a]
Wild type ScLS	88.0	71.0
R40E/H41E-AaLS	69.6	n.d.[b]
R21E/R40E/H41E-AaLS	69.7	n.d.[b]
L121E/I125E-AaLS	57.1	n.d.[b]
R40E/L121E/I125E-AaLS	79.8	n.d.[b]
H41E/L121E/I125E-AaLS	81.5	n.d.[b]
H41R/L121R-AaLS	92.4	73.6
R40E/H41E/L121E-AaLS	86.3	68.6
R40E/H41E/I125E-AaLS	86.0	69.2
R40S/H41S/I125S-AaLS	90.0	71.0
R40S/H41L/L121E-AaLS	90.0	71.3

[a]This protein elutes in the void volume of the S-300 column.

[b]Not determined.

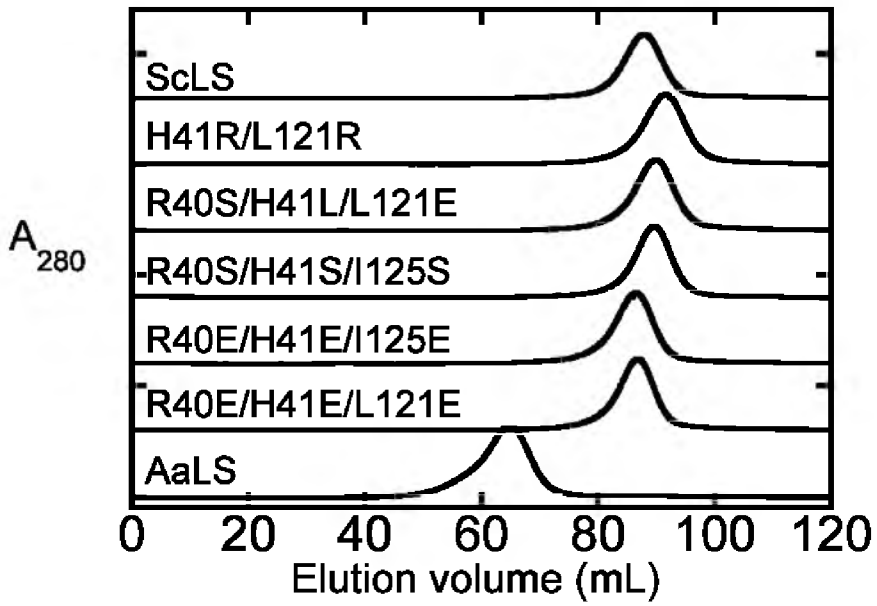


Figure 2.2. SEC elution profiles of the AaLS variants. Proteins were injected onto a sephacryl S-400HR column. AaLS (bottom trace) and ScLS (top trace) are standards representing 60-subunit capsid and pentamer, respectively.

disruption of either the ionic and hydrophobic interactions (R40E/L121E/I125E) or the hydrogen bonding and hydrophobic sites (H41E/L121E/I125E) is sufficient to hinder capsid formation, but these variants still assemble beyond the pentamer stage, perhaps forming dimers-of-pentamers or trimers-of-pentamers. Introducing arginines instead of glutamates at the hydrogen bonding and hydrophobic sites seems to fully prevent any higher order assembly of the pentamers (Figure 2.2). Indeed the H41R/L121R variant might possess the minimal set of mutations sufficient to convert the AaLS capsid to pentamers.

To check for the role of charge repulsion in preventing capsid assembly, we made the R40S/H41S/I125S variant. SEC analysis showed a similar elution profile to the analogous R40E/H41E/I125E variant (Figure 2.2), indicating that the introduction of extra Coulombic repulsion at these positions is not necessary to keep the pentamers from coming together.

The AaLS variants described above were designed without consideration of sequence homology. Nontheless, an amino acid sequence alignment of AaLS with its pentameric homolog from *S. cerevisiae* (ScLS) reveals differences at positions 40, 41, and 121 (Figure 2.3). Therefore, we made the R40S/H41L/L121E variant of AaLS in which the amino acids at these positions correspond to those found in ScLS. This variant and ScLS show similar elution profiles by SEC (Figure 2.2), suggesting that these positions may have been important in the evolution of LS quaternary structure.

The pentameric variants identified above by SEC with the S-400HR column were also injected onto an S-300HR column, which has an optimal size separation range for smaller proteins than the S-400HR column. The elution profiles from the S-300HR column

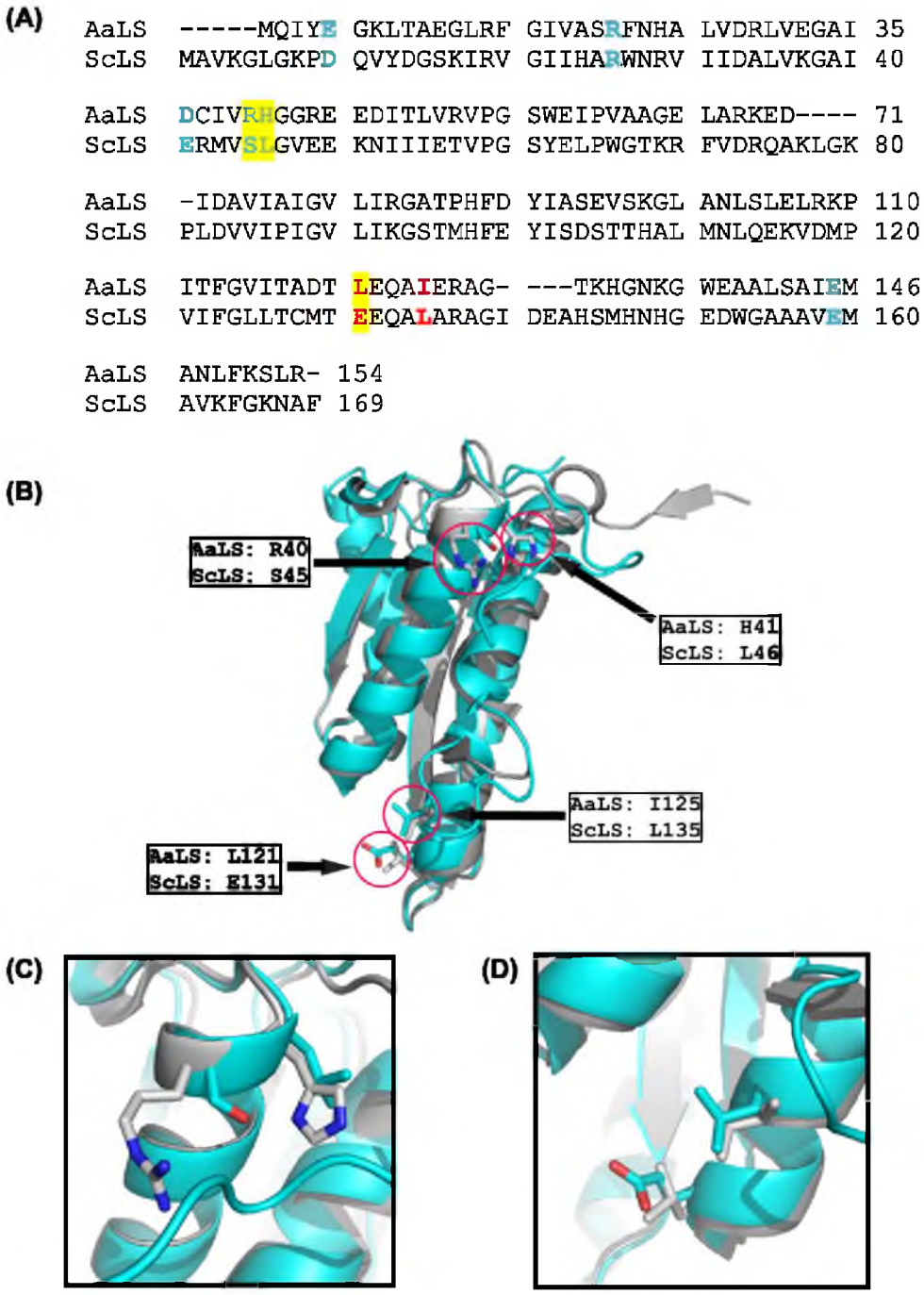


Figure 2.3. Comparison of capsid-forming AaLS and pentameric ScLS. (A) Amino acid sequence alignment of AaLS and ScLS. Positions involved in the ionic, hydrogen bonding, and hydrophobic interaction sites between pentamers in the AaLS capsid are colored blue, green, and red, respectively. Positions 40, 41, and 121 (AaLS numbering) are highlighted in yellow. (B) Superimposition of single subunits of AaLS (gray) and ScLS (cyan). (C) Close-up view of AaLS R40/H41 (ScLS S45/L46) region. (D) Close-up view of AaLS L121/I125 (ScLS E131/L135) region.

exhibited similar trends as the S-400HR column (Table 2.2).

2.2.3. Sedimentation Equilibrium Studies

SEC is a low resolution method for determining quaternary structure. While the results described above unambiguously show a shift to a much lower assembly state for some of the AaLS variants, our designation of the slowest eluting variants as pentamers was somewhat tentative. Therefore, we turned to sedimentation equilibrium experiments in order to confirm the subunit stoichiometries. For all variants, the sedimentation equilibrium data fit well to a single ideal species model and gave predicted masses close to that calculated for the pentamer (Table 2.3, Figures 2.4 and 2.5).

2.2.4. Assessing Folding and Stability of AaLS Variants

To determine whether the dramatic changes in quaternary structure were accompanied by large changes in folding and stability, we utilized CD spectroscopy. LS is a member of the α and β (α/β) class of protein structure. Far-UV CD spectra were obtained to compare the secondary structure content of the AaLS variants. All of the spectra displayed broad minimum between 222 nm and 208 nm, typical of an α/β protein (Figure 2.6A). The similar shapes of the spectra for wild-type AaLS and all of the variants indicates that none of the mutations induced major changes in overall secondary structure. The H41R/L121R-AaLS and R40E/H41E/I125E-AaLS variants had the lowest signal intensities (~12% to 16% less negative θ values than wild-type AaLS between 220 nm and 208 nm), but these differences are likely within the scatter of the experiment (Figure 2.7). Thus, the secondary structure contents of the variants are all similar to wild-type

Table 2.3. Assembly states of AaLS variants determined by sedimentation equilibrium analysis

Protein	Pentameric molecular weight (kDa)[a]	Observed molecular weight (kDa)[b]
wild-type AaLS	N/A[c]	970[c]
wild-type ScLS	98	92
R40S/H41S/ I125S-AaLS	84	79
R40E/H41E/ I125E-AaLS	85	85
R40E/H41E/ L121E-AaLS	85	81
H41R/L121R-AaLS	85	82

[a]Pentameric molecular weights were obtained by calculating the monomeric molecular weight from the amino acid sequence and multiplying by five.

[b]Observed molecular weights were determined by fitting the data from sedimentation equilibrium experiments to a single ideal species model.

[c]Not applicable. Wild-type AaLS forms a 60-subunit capsid with a molecular weight of 1000 kDa.

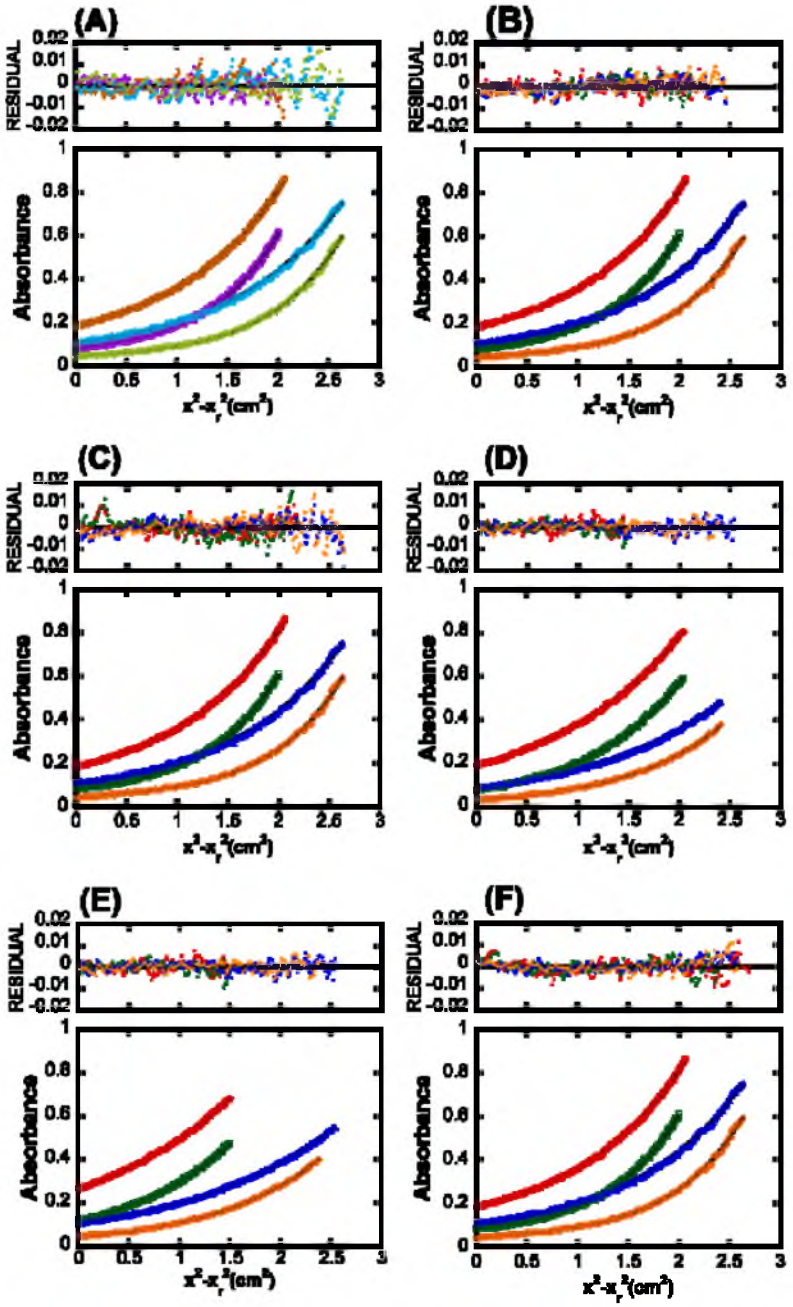


Figure 2.4. Plots of sedimentation equilibrium data. All samples were spun at two different concentrations (0.50 mg/mL and 0.25 mg/mL). (A) Wild type AaLS was spun at 4000 rpm (brown and light blue for 0.50 mg/mL and 0.25 mg/mL samples, respectively) and 5000 rpm (purple and light green for 0.50 mg/mL and 0.25 mg/mL samples, respectively). (B)–(F) Samples were spun at 12500 rpm (red and blue for 0.50 mg/mL and 0.25 mg/mL samples, respectively) and 15000 rpm (green and orange for 0.50 mg/mL and 0.25 mg/mL samples, respectively). (B) Wild-type ScLS. (C) R40E/H41E/L121E-AaLS. (D) R40E/H41E/I125E-AaLS. (E) R40S/H41S/I125S-AaLS. (F) H41R/L121R-AaLS. Black curves represent ideal curves derived from fitting the data to a single ideal species model in all cases. Residuals for each data set are shown above the absorbance plots.

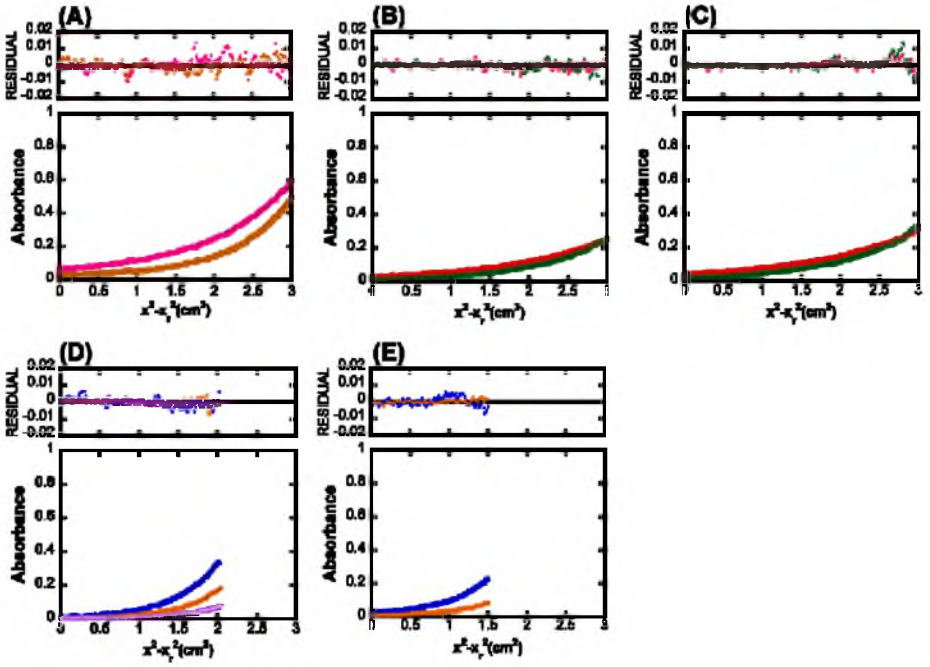


Figure 2.5. Additional sedimentation equilibrium data used in the analysis of some variants. (A) Wild-type AaLS was spun at a concentration of 0.13 mg/mL and rotor speeds of 4000 rpm (pink) and 5000 rpm (brown). (B) ScLS-wt was spun at a concentration of 0.13 mg/mL and rotor speeds of 12500 rpm (red) and 15000 rpm (green). (C) R40E/H41E/L121E-AaLS was spun at a concentration of 0.13 mg/mL and rotor speeds of 12500 rpm (red) and 15000 rpm (green). (D) R40E/H41E/I125E-AaLS was spun at a rotor speed on 20000 rpm and concentrations of 0.5 mg/mL (blue), 0.25 mg/mL (orange), and 0.13 mg/mL (purple). (E) R40S/H41S/I125S-AaLS was spun at a rotor speed on 20000 rpm and concentrations of 0.5 mg/mL (blue) and 0.25 mg/mL (orange). Ideal curves derived from fitting the data to a single ideal species model are shown in black for all samples.

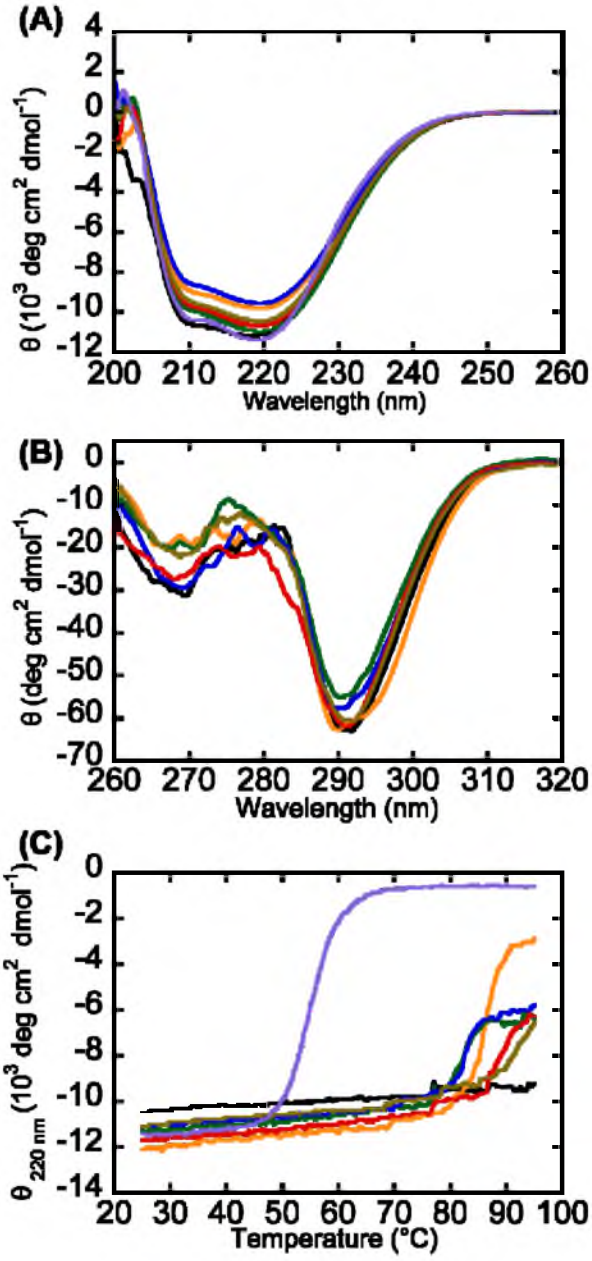


Figure 2.6. Analysis of protein structure and stability by CD spectroscopy. (A) Far-UV CD spectra. (B) Near-UV-CD spectra. (C) Thermal unfolding curves. Proteins analyzed in (A)–(C) are wild-type AaLS (black), ScLS (purple), R40S/H41S/I125S-AaLS (red), R40E/H41E/I125E-AaLS (blue), R40E/H41E/L121E-AaLS (green), R40S/H41L/L121E-AaLS (brown), and H41R/L121R-AaLS (orange).

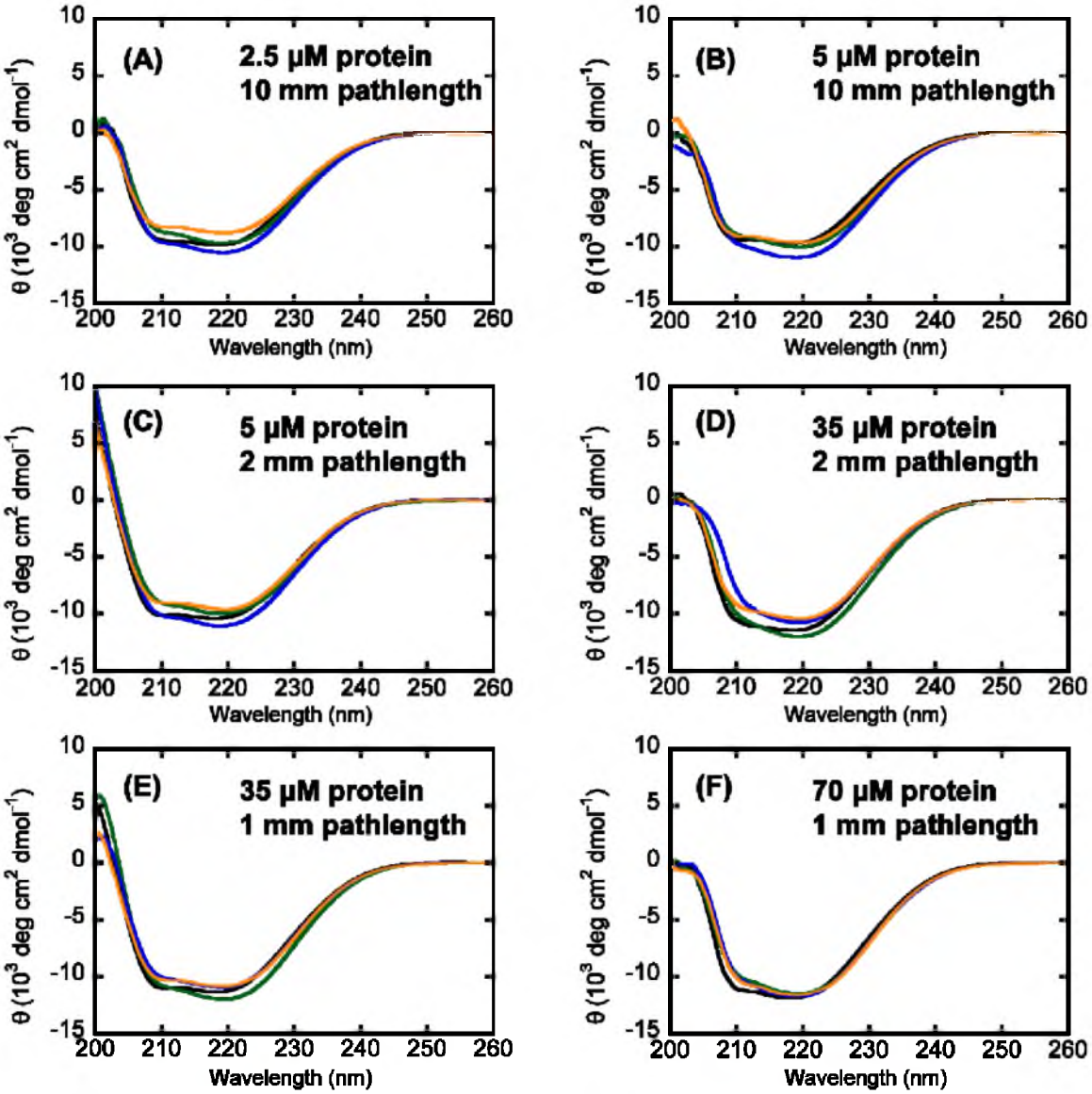


Figure 2.7. Additional far-UV CD spectra. (A) 2.5 µM protein, 10-mm path length. (B) 5 µM protein, 10-mm path length. (C) 5 µM protein, 2-mm path length. (D) 35 µM protein, 2-mm path length. (E) 35 µM protein, 1-mm path length. (F) 70 µM protein, 1-mm path length. Proteins analyzed in (A)–(F) are wild-type AaLS (black), R40E/H41E/I125E-AaLS (blue), R40E/H41E/L121E-AaLS (green), and H41R/L121R-AaLS (orange). Spectra were obtained as described in the Materials and Methods.

AaLS.

The tertiary structures of the AaLS variants were examined by near-UV CD spectroscopy, which reports on the environment around two tyrosines and two tryptophans in each protein. Only minor deviations were observed between the spectra, consistent with preservation of the overall fold (Figure 2.6B). These differences may stem from perturbations around W137, which lies near the pentamer interface in the wild type AaLS capsid.

Wild-type AaLS is exceptionally thermostable with a reported melting temperature of 120 °C.[103] The extent to which the pentamer interfaces contribute to the stability is not known. Therefore, we obtained thermal unfolding curves by monitoring the ellipticity at 220 nm with increasing temperature (Figure 2.6C). Consistent with its reported stability, wild-type AaLS showed little change in signal up to 95 °C. In contrast, the midpoint of the unfolding transition for mesophilic ScLS occurred around 56 °C, which is close to the reported value.[132] All of the pentameric AaLS variants show signs of a partial unfolding transition before 95 °C but remain fully folded up to 80 °C. For R40S/H41S/I125S-AaLS and R40S/H41L/L121E-AaLS, the transitions are incomplete at 95°C. The H41R/L121R-AaLS, R40E/H41E/L121E-AaLS, and R40E/H41E/I125E-AaLS retained a substantial amount of CD signal following the ends of their transitions. The structural basis of this residual signal is not clear, but may indicate a complex unfolding mechanism. To further investigate the thermal denaturation behavior of the proteins, we performed additional measurements in 4 M urea. The far-UV spectra were similar for wild-type AaLS, and the three variants and the partial unfolding transitions of the variants occurred at lower temperatures (Figure 2.8 A&B). Although the magnitudes of the signal changes in the

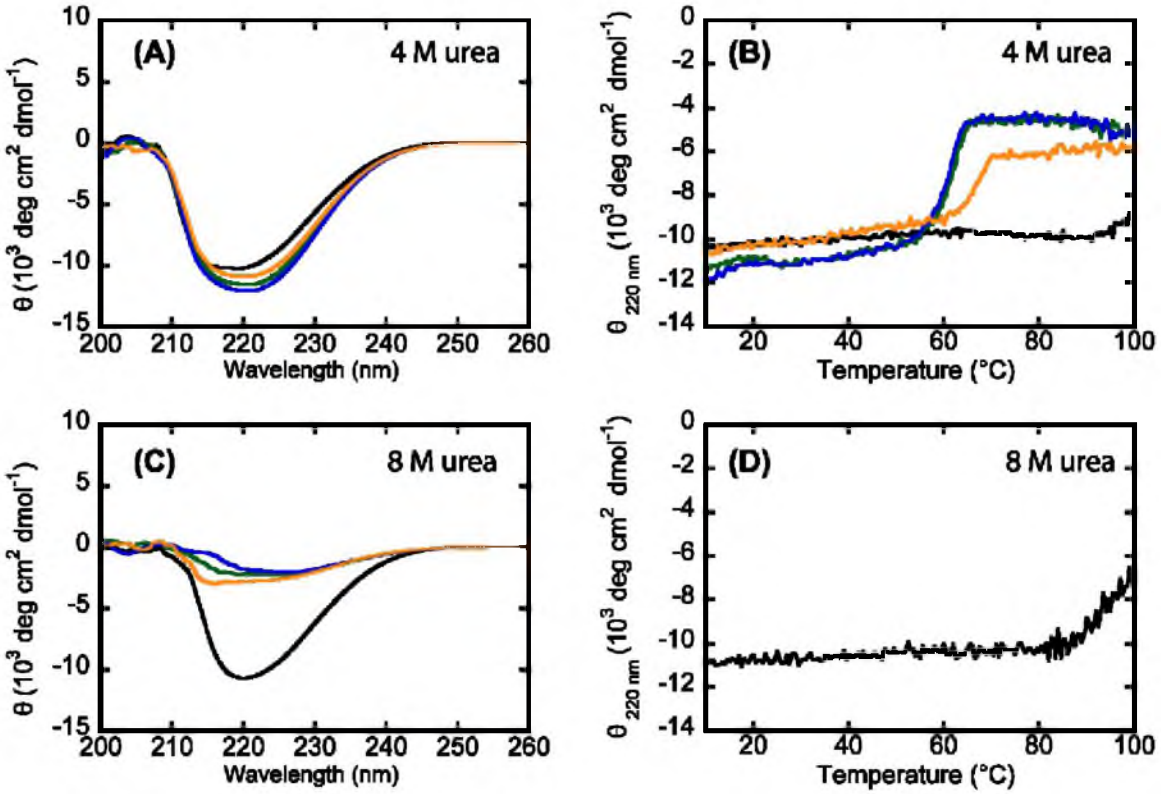


Figure 2.8. CD spectroscopy in the presence of urea. (A) Far-UV CD spectra in pH 8 buffer containing 4 M urea at 10 °C. (B) Melting curves of samples from panel (A). (C) Far-UV CD spectra in pH 8 buffer containing 8 M urea at 10 °C. (D) Melting curve of wild-type AaLS in 8 M urea. For all samples, the path length is 1 mm and the protein concentration is 70 μM. Proteins analyzed in (A)–(C) are wild-type AaLS (black), R40E/H41E/I125E-AaLS (blue), R40E/H41E/L121E-AaLS (green), and H41R/L121R-AaLS (orange).

melting curves are somewhat different in 4 M urea relative to 0 M urea (Figures 2.6C and 2.8B), there is still no evidence of a second transition for these variants up to 99 °C. In contrast to the variants, wild-type AaLS shows very little change in signal between 10 °C and 99 °C in 4 M urea. To further destabilize the proteins, we obtained far-UV CD spectra in 8 M urea. Under this condition, the three variants seem to be fully unfolded at 10 °C, and wild-type AaLS seems to maintain a high degree of structure (Figure 2.8C). Interestingly, wild-type AaLS in 8 M urea does not begin to unfold until ~85 °C, at which point the CD signal intensity slowly and continuously decreases (Figure 2.8D). However, this transition is still not complete at 99 °C. Nonetheless, these data unequivocally show that the loss of the pentamer interfaces in AaLS is accompanied by a significant drop in stability, but that the pentameric variants retain a high degree of thermostability. Further study will be required to understand the unfolding mechanisms of wild-type AaLS and the variants.

2.3.Discussion

The conversion of protein assemblies into independently folded building blocks by targeted disruption of noncovalent interactions at subunit interfaces has a long track record of success. For example, the tetramer of rabbit fructose-1,6-bisphosphate aldolase was converted to either dimers or monomers by one or two point mutations at interface residues .[133,134] Similarly, a single lysine to glutamate mutation was sufficient to convert the normally tetradecameric *E. coli* chaperonin GroEL to monomers.[135] In addition, human arginase has been converted from a trimer to monomers by a glutamate to glutamine substitution that broke an ionic interaction at the three-fold symmetry axis.[136]

The dimeric tyrosyl tRNA synthetase from *Bacillus stearothermophilus* provides another interesting case in which the substitution of aspartate for phenylalanine at the two-fold symmetry axis generated a pH dependent reversible switch for dimer assembly.[137] These examples demonstrate that one or a few mutations that remove noncovalent interactions between subunits (or that introduce new destabilizing interactions at the subunit interface) are often sufficient to change quaternary structure without substantially altering secondary or tertiary structure. However, the loss of protein-protein interfaces is accompanied by a loss of stability in these variants. In the examples above, the symmetry axes of the wild-type complexes proved to be rich sources for designing variants that still fold but do not fully assemble. These principles held true in our pentameric AaLS variants in which both the two- and three-fold symmetry axes of the wild-type capsid were targeted for mutation. These variants also showed signs of lower stability but still folded and retained a great deal of thermostability.

For many large oligomeric proteins, assembly is believed to be a hierarchical process in which monomers first independently assemble to a lower order quaternary state, and these partially assembled building blocks further associate to the final higher order structure.[138] For AaLS, the 60-subunit capsid is believed to assemble from 12 pentameric building blocks.[103] We were able to halt assembly at the pentamer stage by introducing two or three point mutations. The trapping of presumed assembly intermediates by rationally designed point mutations has also been accomplished for other proteins. For example, inorganic pyrophosphatase from *E. coli* was converted from a hexamer to trimers by simultaneously mutating two histidines to glutamines at the two-fold symmetric interface of the hexamer.[139] In another case, two mutations (tryptophan to

alanine and methionine to alanine) were used to isolate the hexameric form of HIV capsid protein.[140] The inability of these hexamers to further assemble was ascribed to the removal of a hydrophobic interaction at the two-fold symmetry axis between hexameric building blocks of the wild-type capsid. Of particular relevance to our study, the assembly states of both maxi- and miniferritins were simplified by minimal point mutations. For maxiferritin (bacterioferritin from *E. coli*), the 24-subunit capsid was converted to folded dimers by several different single amino acid substitutions at the two-fold and three-fold symmetry axes.[141] For miniferritin (*E. coli* DNA binding protein from starved cells), two mutations at the three-fold symmetry axis led to dimers rather than 12-subunit capsids.[142] In these studies, wild-type residues were replaced by alanines to remove stabilizing electrostatic interactions at subunit interfaces. Thus, the strategy of simplifying complex quaternary structures by mutating residues at symmetry axes between building blocks may be broadly applicable.

A definitive understanding of assembly state determinants in LSs has remained elusive. Previous attempts to design disruptions in the interactions between pentamers of capsid-forming LSs gave rise to interesting changes in capsid structure rather than preventing further assembly of pentamers.[107] By visual inspection of the AaLS capsid structure, we identified a set of amino acid side chains that formed stabilizing interactions between pentamers.

By mutating subsets of these residues, we induced quaternary structure changes in AaLS. Simultaneous disruption of the ionic, hydrogen bonding, and hydrophobic interaction sites reliably yields stable pentamers. Interestingly, introduction of negatively charged residues at the hydrophobic site alone led to an expansion of the capsid.

Presumably the larger capsid size helps to minimize Coulombic repulsion in the variant, perhaps by increasing the size of the pore at the three-fold symmetry axis. Since the stabilizing effects of individual protein-protein interactions are multiplied by the symmetry of the capsid,[1] it is perhaps unsurprising that the interfaces are highly plastic and resilient to localized disruptions. Indeed, a similar capsid expansion was found to relieve steric clashes at the pentamer-pentamer interface in AaLS upon the insertion of a four amino acid loop between amino acids 129 and 130 (*vide infra*). Mutation of only two of the three interaction sites was often sufficient to break the capsid but did not always completely abolish associations between pentamers. Although a more detailed characterization would be required to assign the precise assembly states of the R40E/L121E/I125E-AaLS and H41E/L121E/I125E-AaLS variants, SEC analysis indicates that they form either dimers-of-pentamers or trimers-of-pentamers. The lumazine synthase from *Brucella abortus* (BaLS) provides a natural precedent for a pentamer dimer, in which ten total subunits assemble into a D_5-symmetric complex. The pentamer-pentamer interface of BaLS involves π-stacking between F121 of one pentamer and H124 and F128 of the other pentamer (using the numbering of Ainciart *et al.*).[129] Intriguingly, these residues correspond to L121, R127, and an insertion proximal to K131 of AaLS, respectively, based on their amino acid sequence alignment. It is tempting to speculate that a pentamer dimer of AaLS variants might arise from a salt bridge between the newly introduced glutamate at position 121 and the positively charged residues at positions 127 and 131. A D_5-symmetric assembly would lead to the formation of 10 such salt bridges (one per subunit) distributed over this new interface, which could provide substantial stabilization. It is difficult to envision what a trimer-of-pentamers might look

like, given the requirements of symmetric assembly;[1] any arrangement of three pentamers is likely to possess open sites for the binding of more pentamers via interactions equivalent to those in the initial trimer of pentamers. However, it is worth noting that a variant of ScLS, in which three arginines were introduced at the five-fold symmetry axis, has been reported to undergo a dynamic exchange between a pentamer and a trimer-of-pentamers.[132] Altogether, the quaternary structures of the variants reported here implicate residues R40, H41, L121, and I125 as important contributors to capsid assembly.

It should be noted that these sequence positions may not be the sole determinants of AaLS capsid assembly. For example, R21 may play a comparable role to R40 in stabilizing the capsid, as part of the ionic network (Figure 2.1D). Additional residues that we did not consider in our designs might also act as determinants of capsid assembly.

A previous study by Fornasari *et al.*[130] used sequence and structural analysis to identify a set of eight amino acids (G6, N23, D36, I125, E126, G129, K131, and E145) responsible for capsid formation by LSs. Interestingly, only one of these eight sequence positions (I125) was changed in our pentameric variants, and some of our pentameric variants did not alter any of these eight sequence positions. While this previous study did not include AaLS, its homology with other capsid forming LSs suggests that our findings are relevant across this family. Of the seven positions that we did not mutate, our analysis of the noncovalent interactions between pentamers in wild type AaLS also identified D36 and E145 as part of an ionic network (Figure 2.1D). The main chain of G6 contributes a hydrogen bond to the ionic network, which may also be significant for capsid assembly. E126 lies next to L121' on the neighboring pentamer, and a contact between C_γ of E126 and a C_δ of L121' may help stabilize the hydrophobic cluster. G129 makes a main-chain–

main-chain hydrogen bond with A25'; mutation of this amino acid could also potentially hinder association of pentamers by disrupting this favorable interaction and/or introducing unfavorable steric clashes. N23 lies in close proximity to an adjacent pentamer, and its side chain carbonyl oxygen may form a stabilizing $n \rightarrow \pi^*$ interaction[143] with the backbone carbonyl carbon of T130'. In the AaLS crystal structure, K131 does not appear to make close contacts with any other residue and projects to the capsid interior.

The first 10 N-terminal residues have also been proposed to act as determinants of capsid formation,[99] based on whether this region can extend a β-sheet from an adjacent subunit within the same pentamer. Our findings suggest an alternative role for the N-terminus in stabilizing the ionic network through proper positioning of E5 and G6. A kink in the C-terminal α-helix with a sequence motif of G(T/G)K(A/H)GN at positions 129 to 134 has also been identified as a likely determinant of capsid assembly.[98,130] This motif includes G129 and K131 from the set of eight amino acids discussed above. Further, pentameric LSs often have insertions of 1 to 4 amino acids following G129, which likely hinder capsid formation by introducing steric clashes at the pentamer-pentamer interfaces.[97,107,130] However, insertion of four amino acids into AaLS at this position does not prevent capsid assembly.[107] The expanded structures formed by this insertion variant underscore the plasticity of the protein-protein interfaces in the capsid. Our results suggest that it is not necessary to perturb the structure of the protein in this region in order to prevent capsid assembly. While these other proposed capsid determinants are not inconsistent with our findings, identifying and verifying additional sets of residues that can convert the capsid into pentamers will require further study.

In addition to advancing a basic understanding of LS capsid assembly, our results may also aid in efforts to utilize this scaffold for bionanotechnology. AaLS has shown much promise as a customizable molecular container.[111,112] The ability to control assembly state would potentially improve the versatility of AaLS based encapsulation systems. The pentameric variants described herein might provide valuable jumping off points for the design of reversible capsid assembly switches. Controlled assembly and disassembly of the AaLS capsid would enable the loading of cargo molecules *in vitro*, which could help make AaLS a powerful tool for applications such as drug delivery, catalysis, or the synthesis of nanometerials.

2.4. Materials and Methods

2.4.1. Materials

Cell culture media, salts, and chemical reagents were purchased from Sigma-Aldrich, Fisher scientific, Pierce biotechnology, or Bio-rad (unless otherwise noted) and used without further purification. *Pfu*-turbo DNA polymerase and *E. coli* strains BL21 (DE3) and XL1-Blue were purchased from Stratagene. DpnI endonuclease was purchased from New England Biolabs. The oligonucleotides used in this study were synthesized by the DNA/peptide synthesis Core Facility at the University of Utah.

2.4.2. Site-directed Mutagenesis

Plasmid pMG-AaLS,[111] encoding the wild-type AaLS gene with an eight amino acid extension (LEHHHHHH) at the C-terminus, was a kind gift from Prof. Donald Hilvert (ETH Zürich, Zurich, Switzerland). To avoid potential assembly artifacts, we removed

the His-tag by mutating the first two histidine codons to stop codons. Using pMG-AaLS plasmid as the template DNA, site-directed mutagenesis PCR was carried out with mutagenic primers HCN61 and HCN62 (Table 2.4). The PCR reaction mixture (50 μL) was prepared by combining 35 μL of water, 5 μL of 10x reaction buffer, 1 μL of *Pfu*-turbo DNA polymerase (2.5 U/μL), 5 μL of dNTPs (2 mM each), 1 μL of template DNA (10 ng/μL), and 1.5 μL each of forward and reverse primers (10 μM each). During the mutagenesis reaction, temperature cycling was performed in a Mastercycler personal thermocycler (Eppendorf). The *Pfu*-turbo DNA polymerase was activated at 95 °C for 1 min before starting the temperature cycling. The mutagenesis reaction involved 20 cycles of denaturation (95 °C for 30 s), annealing (55 °C for 30 s), and primer extension (72 °C for 6 min). Following the last cycle, the reaction was incubated for an additional 10 min to let the DNA polymerase finish the final elongation before cooling the mixture to 37 °C. The restriction endonuclease DpnI (20 U) was then added to the reaction, which was incubated for an additional 60 min to remove the original DNA template. The mutagenesis reaction product was transformed into $CaCl_2$-competent XL1-Blue cells by heat shock (42 °C for 1 min). The transformed cells were plated onto LB-agar containing ampicillin (100 μg/mL) and incubated overnight at 37 °C. The resulting colonies were used to obtain stocks of plasmid DNA with a QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. The coding portion of the resulting plasmid (pMG-AaLSNoHis) was confirmed by DNA sequencing (University of Utah DNA sequencing core facility).

Additional mutations in the coding region of pMG-AaLSNoHis were also generated by the procedure described above. For these mutagenesis reactions, pMG-AaLsNoHis

Table 2.4. Oligonucleotide primers used for site-directed mutagenesis. Sequences are given in the 5′ to 3′ direction. Mutagenic bases are highlighted in bold.

Primer Name	DNA sequence (5′ → 3′)
HCN31	ATTACAGCTGACACC**GA**GGAACAGGCT**GAG**GAGCGCGCCGGCAC
HCN32	GTGCCGGCGCGCTC**CTC**AGCCTGTTCC**TC**GGTGTCAGCTGTAAT
HCN55	CGTTTCGGTATCGTAGCATCA**GAG**TTTAATCATGCTCTTGTCGAC
HCN56	GTCGACAAGAGCATGATTAAA**CTC**TGATGCTACGATACCGAAACG
HCN57	GGTGCAATTGATTGCATAGTC**GAG**CATGGCGGCCGTGAAGAAGAC
HCN58	GTCTTCTTCACGGCCGCCATG**CTC**GACTATGCAATCAATTGCACC
HCN61	CTTATTCAAGTCTCTCCGACTCGAG**TA**A**T**A**G**CACCACCACCACTAAAT AATG
HCN62	CATTATTTAGTGGTGGTGGTG**C**T**AT**T**A**CTCGAGTCGGAGAGACTTGAA TAAG
HCN65	GCAATTGATTGCATAGTCCGT**GA**G**G**GCGGCCGTGAAGAAGAC
HCN66	GTCTTCTTCACGGCCGCC**C**T**C**ACGGACTATGCAATCAATTGC
HCN67	GAGGGTGCAATTGATTGCATAGTC**GAGGA**G**G**GCGGCCGTGAAGAAGAC
HCN68	GTCTTCTTCACGGCCGCC**C**T**CCTC**GACTATGCAATCAATTGCACCCTC
HCN73	GTTATTACAGCTGACACC**GA**GGAACAGGCTATCGAGCGC
HCN74	GCGCTCGATAGCCTGTTCC**TC**GGTGTCAGCTGTAATAAC
HCN75	GACACCTTGGAACAGGCT**GAG**GAGCGCGCCGGCACAAAAC
HCN76	GTTTTGTGCCGGCGCGCTC**CTC**AGCCTGTTCCAAGGTGTC
HCN97	GAGGGTGCAATTGATTGCATAGTC**AGCAGC**GGCGGCCGTGAAGAAGAC
HCN98	GTCTTCTTCACGGCCGCC**GCTGCT**GACTATGCAATCAATTGCACCCTC
HCN99	CCTTGGAACAGGCTA**G**CGAGCGCGCCGGCAC
HCN100	GTGCCGGCGCGCTCG**C**TAGCCTGTTCCAAGG
HCN107	CAATTGATTGCATAGTCAGC**TTG**GGCGGCCGTGAAGAAGAC
HCN108	GTCTTCTTCACGGCCGCC**CAA**GCTGACTATGCAATCAATTG

served as the DNA template. The mutagenic primers used in these reactions are listed in Table 2.4.

2.4.3. Protein Production and Purification

For producing AaLS variant proteins, $CaCl_2$ competent BL21 (DE3) *E. coli* cells were transformed with the pMG-AaLSNoHis plasmid containing the desired mutations. The cells were cultured in 500 mL of LB medium with ampicillin (100 mg/L) at 37 °C and 250 rpm until the $O.D._{600}$ reached 0.7. At this point, protein overproduction was induced by the addition of IPTG (0.1 mM final concentration). Following incubation for an additional 20 h at 30 °C and 250 rpm, the cells were harvested by centrifugation (5000 x *g* for 10 min at 4 °C). The cell pellets were resuspended in 10 mL lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0). The cells were lysed by incubation with lysozyme (5 mg), RNase A (0.6 mg), and DNase I (10 U) on ice for 30 min followed by sonication. The cell lysate was cleared by centrifugation (8000 x *g* for 30 min at 4 °C). The cleared lysate (10 mL) was then combined with a saturated ammonium sulfate solution (5 mL), and this mixture was incubated at room temperature for 30 min. Precipitated impurities were removed by centrifugation (8000 x *g* for 30 min at 4 °C). The supernatant was then dialyzed at 4 °C in ion-exchange buffer A (50 mM sodium phosphate, 20 mM NaCl, pH 8.0) and then loaded onto an ÄKTA FPLC (GE Healthcare) equipped with an anion-exchange column (MonoQ 5/50 GL, GE Healthcare) that was equilibrated with ion-exchange buffer A. Protein was eluted with a gradient of 0–100% ion exchange buffer B (50 mM sodium phosphate, 1 M NaCl, pH 8.0) over 40 mL at 4 °C. The purity of protein was analyzed by SDS-PAGE. The yields of purified proteins ranged

from 30 to 40 mg protein per L of cell culture.

2.4.4. Size-exclusion Chromatography

The approximate sizes of the AaLS variants were analyzed by SEC using an ÄKTA FPLC system equipped with HiPrep 16/60 Sephacryl S-300 HR and HiPrep 16/60 Sephacryl S-400 HR columns (GE Healthcare). The running buffer contained sodium phosphate (50 mM), NaCl (200 mM), and EDTA (5 mM) at pH 8.0. The flow rate was 0.5 mL/min, and the running temperature was 4 °C. Assembly states of the variants were determined based on comparison to wild-type AaLS (an 1100 kDa, 60-subunit capsid), *Saccharomyces cerevisae* LS (a 98 kDa pentamer),[132] and AaLS-neg (a 3300 kDa, 180-subunit capsid).[111]

In all cases, the protein purified using the procedure described above gave two peaks: one containing AaLS variant and one containing impurity. The peak containing impurity always eluted late relative to the AaLS variant. This impurity is not stained by Coomassie blue and displayed a high A_{260}/A_{280} ratio. Therefore, the impurity peak probably does not contain protein, but may be a nucleic acid contaminant. To check whether this contaminant might influence assembly, proteins were reinjected onto the HiPrep 16/60 S-400 HR column. The position and shape of the protein peak was unchanged in all cases. All further analyses of the AaLS variants used protein that had been run over the HiPrep 16/60 S-400 HR column to remove the nucleic acid contaminant.

2.4.5. Sedimentation Equilibrium Analysis

Sedimentation equilibrium experiments were performed on Beckman XL-A analytical ultracentrifuge equipped with an An60Ti rotor and photoelectric scanner. The protein samples were dialyzed into buffer containing sodium phosphate (50 mM), NaCl (200 mM), and EDTA (5 mM) at pH 8.0. Three 160 µl protein samples (130–500 µg/mL) were separately loaded onto three channels of a six-channel centerpiece, and the other channels were loaded with buffer as references. Protein samples were spun at speeds ranging from 12500 to 20000 rpm (except for wild-type AaLS, which was spun at 4000 rpm and 5000 rpm) at 20 °C. Upon reaching equilibrium at a given speed, the UV absorbance at 280 nm was measured with a step size of 0.001 cm and 10 averages. The obtained data were analyzed using Ultrascan (University of Texas Health Science Center) and fit to a single ideal species model (Equation 2.1), where A is the absorbance at radius x, A_0 is the absorbance at a reference radius x_0, M is the molecular weight of a single species, $\bar{v}$ is the partial specific volume of the protein, ρ is the density of the buffer, ω is the angular velocity of the rotor, R is the universal gas constant, T is the absolute temperature, and B is the baseline correction factor. The values of $\bar{v}$ and the ρ at 20 °C were calculated using the program Sednterp.

$$A = \exp\left[(\ln(A_0) + M\omega^2(1-\bar{v}\rho)(x^2 - x_0^2))/2RT\right] + B \qquad (2.1)$$

2.4.6. Circular Dichroism Spectroscopy

The CD spectroscopy experiments were performed on a Jasco J-815 spectropolarimeter. The protein samples were dialyzed into buffer containing sodium

phosphate (50 mM) and NaCl (200 mM) at pH 8.0 at 4 °C. Far-UV CD spectra of AaLS variants (4 μM) were measured in a 1 cm pathlength cuvette at 25 °C by averaging 10 wavelength scans from 200 to 260 nm in 0.1 nm steps with a signal averaging time of 1 s and a bandwidth of 1 nm. Near-UV CD spectra of AaLS variants (160 μM) were measured in a 1 cm pathlength cuvette at 25 °C by averaging 10 wavelength scans from 260 to 320 nm in 0.1 nm steps with a signal averaging time of 1 s and a bandwidth of 1 nm. Thermal denaturation curves of AaLS variants (5 μM) were obtained by measuring the ellipticity at 220 nm in a 1-cm pathlength cuvette from 25 to 95 °C in 0.5 °C intervals with a scan rate of 0.1 °C/min and a signal averaging time of 1 s.

CHAPTER 3

AN ENGINEERED REDOX SWITCH FOR PROTEIN CAPSID ASSEMBLY

3.1. Introduction

The majority of proteins self-assemble into symmetric architectures.[1] However, protein assembly processes are not generally well understood. Protein-protein interfaces tend to be large and chemically complex, involving multiple types of noncovalent interactions.[13] Further, the stability determinants of a protein-protein interaction tend to be unevenly distributed, with a few "hot spot" residues contributing a disproportional share of the binding energy.[25,141] For complexes containing multiple subunits, the assembly mechanism can also be quite complicated.[38] Nonetheless, detailed knowledge of the interplay between structure and energetics in protein-protein interactions can be used to control construction of protein architectures.[115,144,145] Such tailor-made assemblies could provide useful tools for protein biochemistry and nanotechnology.

Many proteins self-assemble into polyhedral capsids that enclose a hollow interior space. Such structures can act as molecular containers. In nature, protein capsids function as hosts for a variety of guests, including nucleic acids (viruses),[146] metals (ferritin),[57] and other proteins (multi-enzyme complexes).[34,35,42,46] Consequently, there has been recent interest in exploiting protein capsids for applications in catalysis, nanomaterials, bioimaging, and drug delivery.[60,67,76,118,147]

Some capsids possess inherent assembly switches, which can facilitate the loading of cargo molecules. For example, the cowpea chlorotic mottle virus (CCMV) capsid protein assembles into a 180-subunit icosahedron at pH 5.0 but disassembles into dimers at pH 7.5.[41,148] This assembly switch has been exploited to develop novel protein encapsulation systems.[80,83] In a few cases, it has been possible to engineer metal-dependent assembly switches for the assembly of polyhedral protein complexes. To increase the useful pH

range of CCMV-based encapsulation systems, a hexahistidine sequence was fused to the N-terminus of the CCMV capsid protein. Addition of nickel ions promotes capsid formation by this variant at pH 7.5.[89] More recently, a metal binding site was installed at a protein-protein interface in ferritin, which acted as a template for copper-dependent assembly.[88] Despite these successes, general strategies for installing assembly switches into proteins are still needed.

Lumazine synthase from *Aquifex aeolicus* (AaLS) is a promising scaffold for developing novel molecular encapsulation systems. Upon self-assembly of 60 identical subunits *in vivo*, this protein forms a dodecahedral capsid with 532 symmetry that can be viewed as a dodecamer-of-pentamers.[103] However, methods are lacking for capsid assembly *in vitro*. The AaLS capsid can be produced efficiently in *Escherichia coli* cells, which has enabled the development of encapsulation systems for both protein and RNA cargoes.[111–113] These *in vivo* encapsulation systems are driven by engineered charge complementarity and require coproduction of the capsids and guests in the same cell, where guest loading presumably occurs during capsid assembly. The AaLS homolog from *Bacillus subtilis* has been used to encapsulate Fe(II) ions *in vitro* because this guest is small enough to diffuse freely through pores in the capsid shell.[110] However, the largest pores in AaLS are ~9 Å wide,[103] which restricts the set of potential guest molecules that can be loaded into AaLS capsids *in vitro*.

The ability to assemble AaLS capsids in a controlled fashion could expand the opportunities for guest loading *in vitro*, enhancing the versatility of this scaffold for molecular encapsulation. Although wild-type AaLS lacks an *in vitro* capsid assembly switch, much is known about the structural determinants for AaLS capsid assembly. In

particular, three distinct interaction sites (ionic network, hydrogen bonding, and hydrophobic interactions) between the pentameric building blocks have been shown to play important roles in stabilizing the capsid.[149] Among the three interaction sites, the ionic network and hydrogen bonding sites are located around the two-fold symmetry axis while the hydrophobic interaction site sits at the three-fold symmetry axis (Figure 3.1). The ionic network consists of two arginine residues (R21 and R40', from different subunits) that are surrounded by three acidic residues (E5, E145, and D36'). The hydrogen bonding interaction is formed by two symmetry-related histidine residues, H41 and H41', from two adjacent subunits. The hydrophobic interaction site involves six residues, three symmetry-related leucines (at position 121) and three symmetry-related isoleucines (at position 125). Simultaneous substitutions of R40, H41, and either L121 or I125 can often be sufficient to halt AaLS assembly at the pentameric stage. Thus, these three residues are involved in interactions that are crucial for capsid formation by AaLS.

3.2. Results

3.2.1. Design of a Redox Switch for Capsid Assembly *In Vitro*

We reasoned that covalent modification of a pentameric AaLS variant could trigger capsid assembly by recapitulating the hydrophobic interaction site from the wild-type AaLS structure. Position 125 represents a promising place to introduce such a modification. Previously, we had found that R40E/H41E/I125E-AaLS forms a stable pentamer but that R40E/H41E-AaLS assembles into capsid. Thus, we can control the self-assembly of AaLS by changing the polarity at position 125.

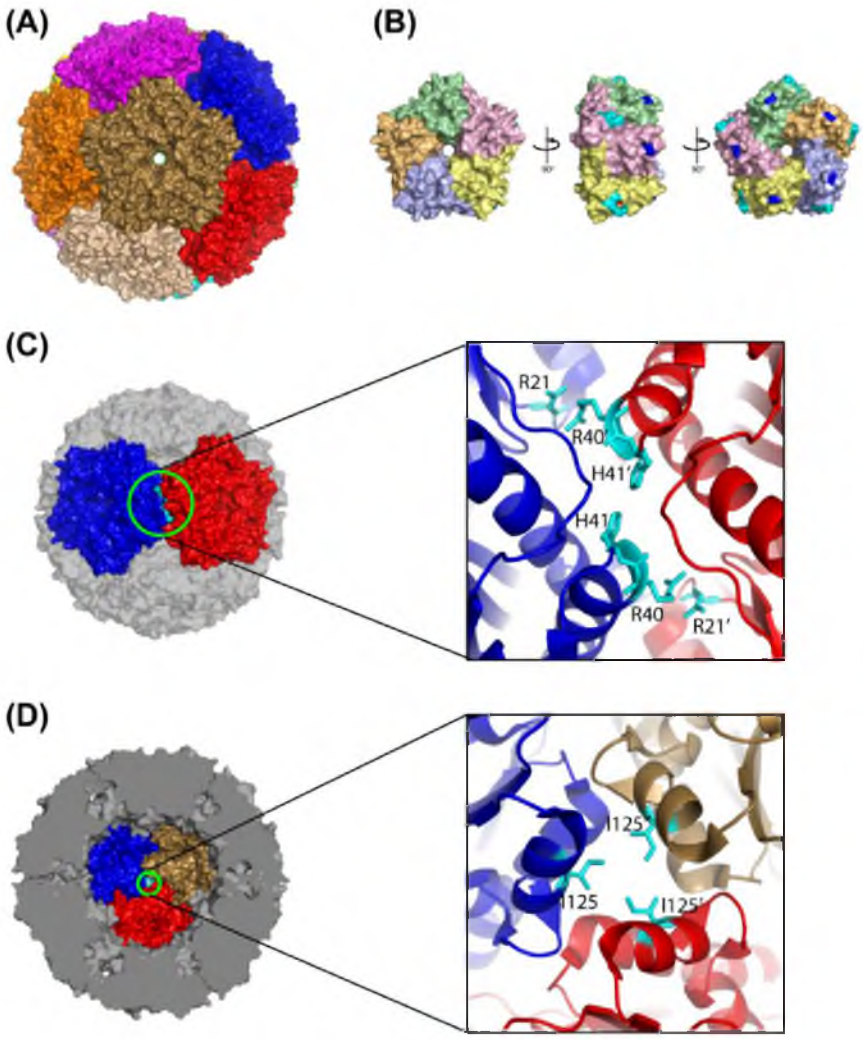
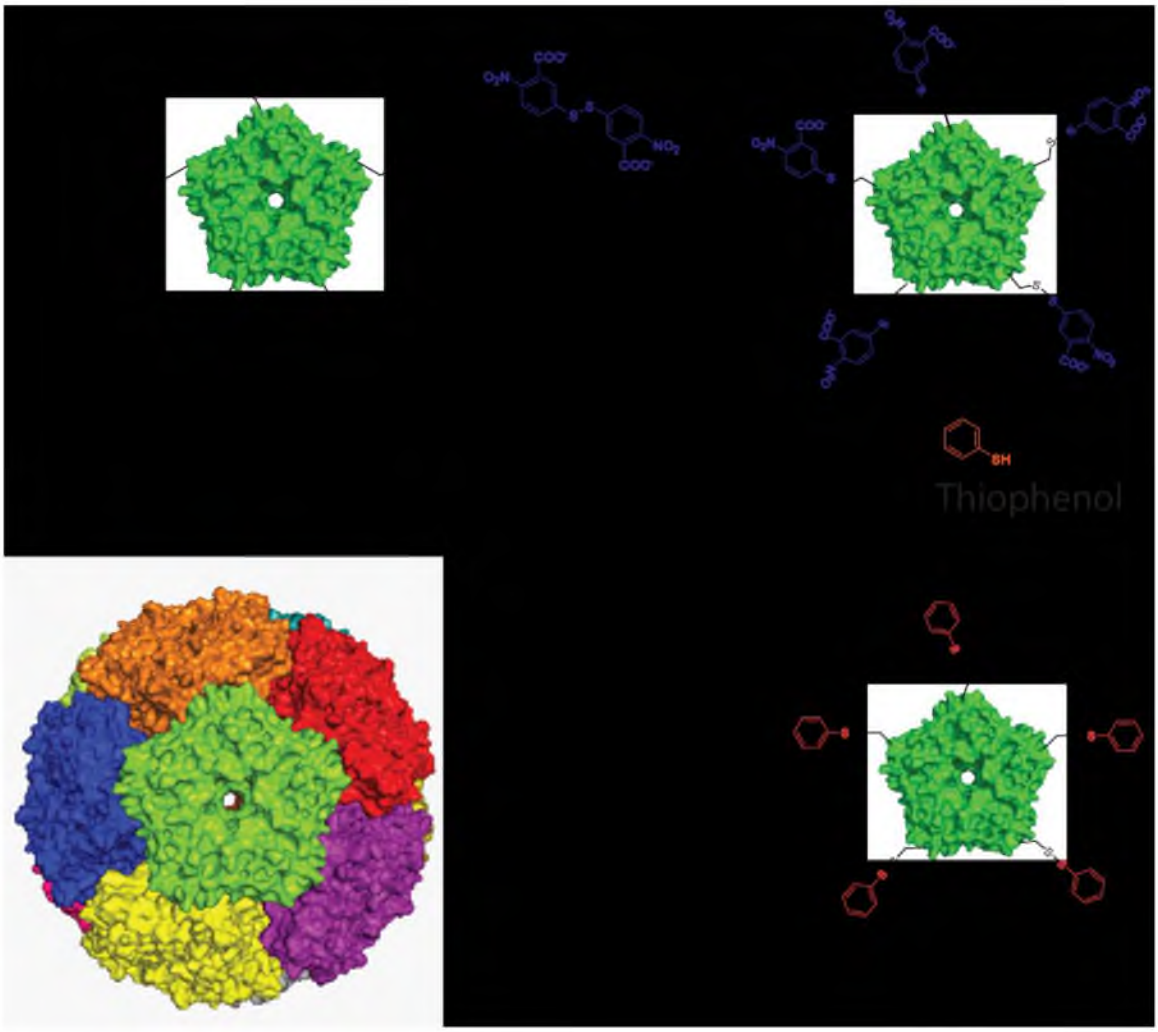


Figure 3.1. Overall structure and interactions between pentamers of the wild-type AaLS capsid. (A) Surface representation of the dodecahedral capsid formed by wild-type AaLS (PDB id: 1HQK). Each pentameric building block is colored differently, although the capsid is composed of 60 identical subunits. (B) Different views of a single pentamer excised from the wild-type AaLS capsid structure. Each perspective represents a 90° rotation relative to its neighbor, as indicated. Each subunit in the pentamer is colored differently for clarity. Amino acids near the two-fold symmetry axis are colored cyan and amino acids near the three-fold symmetry axis are colored dark blue. (C) The wild-type AaLS capsid exterior viewed down a two-fold symmetry axis. Left: A surface representation of the capsid with two adjacent pentamers colored in red and blue (all other subunits are grey). Residues at the pentamer-pentamer interface are shown in cyan. Right: Zoomed-in view of the two-fold symmetric interface between adjacent pentamers. (D) A cross-section of wild-type AaLS showing the capsid interior viewed down a three-fold symmetry axis. Left: a surface representation of the capsid with three adjacent pentamers colored in red, brown, and blue (all other subunits are grey). Residues at the pentamer-pentamer interfaces are shown in cyan. Right: Zoomed-in view of the three-fold symmetric interface between adjacent pentamers.

3.2.2. Capsid Assembly Mediated by Thiophenol Modification

The introduction of a cysteine at this position affords the opportunity to change the hydrophobicity of this residue by thiol-disulfide exchange reactions (Scheme 3.1). To test this hypothesis, we constructed C37A/R40S/H41S/I125C-AaLS (AaLS-switch-red), which was produced in BL21 (DE3) *E. coli* cells and purified by an established method. As expected, AaLS-switch-red formed pentamers (Figure 3.2, blue chromatogram). To trigger capsid formation, AaLS-switch-red was covalently modified with thiophenol via a two-step thiol-disulfide exchange process. In the first step, the protein was oxidized with 5,5'-dithiobis-(2-nitrobenzoate), which generated a disulfide bond between 2-nitro-5-thiobenzoate (NTB) and the cysteine at position 125 of AaLS-switch-red. In the second step, the thiophenol adduct (AaLS-switch-ox) was formed by displacement of the NTB group. Both thiol-disulfide exchange reactions gave nearly quantitative yields as assessed by UV-Vis spectroscopy (which monitored the release of NTB at 412 nm) and mass spectrometry (Table 3.1). The initial characterization of the AaLS-switch-ox assembly state by size-exclusion chromatography (SEC) indicated that the modified protein did indeed show a tendency to further assemble, albeit inefficiently. Three weeks after modification with thiophenol, AaLS-switch-ox formed a mixture of pentamers and some higher order assemblies that were sized in between pentamers and 60-subunit capsids (Figure 3.3A). However, upon addition of PEG-3350, a macromolecular crowding agent known to promote protein association,[150] freshly prepared AaLS-switch-ox gave a high yield of 60-subunit capsids within 4 days (Figure 3.2, green chromatogram). In contrast, incubation of AaLS-switch-red pentamers with PEG-3350 failed to produce capsids (Figure 3.3B), showing that thiophenol modification is necessary for capsid formation.



Scheme 3.1. Scheme for the *in vitro* assembly of the AaLS capsid via thiophenol modification. The upper left structure represents an AaLS-switch-red pentamer. The bottom right structure represents an AaLS-switch-ox pentamer. The bottom left structure represents an AaLS-switch-ox capsid constructed from 12 identical, but differently colored, AaLS-switch-ox pentamers.

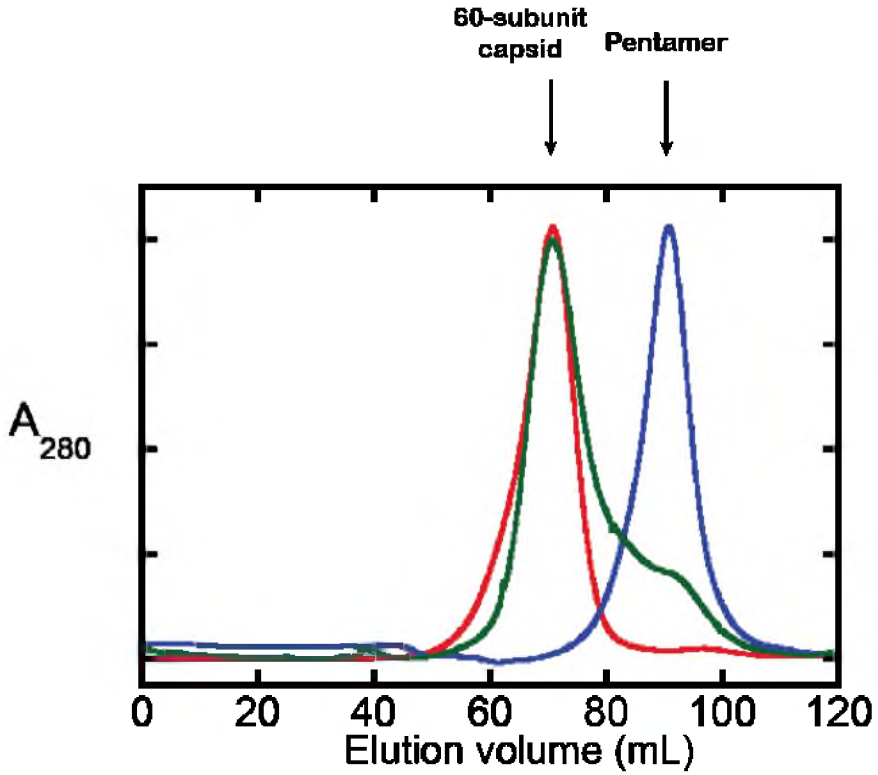


Figure 3.2. Assembly of AaLS-switch-ox *in vitro*. Size-exclusion chromatograms for wild-type AaLS (red), AaLS-switch-red (blue), and AaLS-switch-ox (green). Protein samples were eluted from a HiPrep 16/60 Sephacryl S-400 HR column at 4 °C and pH 8.0.

Table 3.1. Mass spectrometry characterization of AaLS-switch-red, DTNB treated AaLS-switch-red, and AaLS-switch-ox.

Protein	Calculated mass (Da)	Observed mass (Da)
AaLS-switch-red (no additional modification	16787	16787
AaLS-switch-NTB conjugate (NTB-protein adduct)	16985	16983
AaLS-switch-ox (thiophenol-protein adduct)	16896	16895

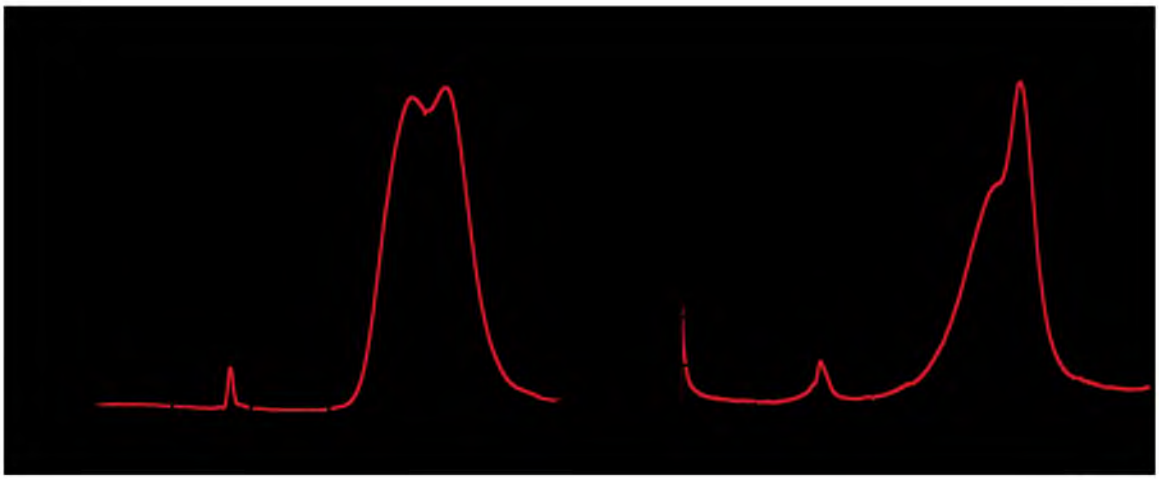


Figure 3.3. Control experiments for assembly of AaLS-switch-ox *in vitro*. (A) Size-exclusion chromatogram, obtained with a HiPrep 16/60 Sephacryl S-400 HR column, of AaLS-switch-ox that was not treated with PEG-3350. Following modification with thiophenol, the protein sample was incubated in SEC running buffer (50 mM sodium phosphate and 200 mM NaCl, pH 8.0) at 4 °C for 3 weeks before injection onto the column. (B) Size-exclusion chromatogram, obtained with a HiPrep 16/60 Sephacryl S-400 HR column, of AaLS-switch-red incubated for 2 days at room temperature in SEC running buffer containing 10 % (w/v) PEG-3350 and 10 mM DTT.

While PEG-3350 was also necessary for capsid formation, the capsids were stable for at least 3 months after its removal by SEC. All further experiments on AaLS-switch-ox were done with SEC-purified capsids.

To confirm the proper capsid assembly by AaLS-switch-ox, individual particles were visualized by transmission electron microscopy. AaLS-switch-ox forms uniform assemblies that closely resemble wild-type AaLS in both size and shape (Figure 3.4). The average diameter for wild-type AaLS is 17 ± 1.7 nm (307 particles). The average diameter for AaLS-switch-ox capsid is 18 ± 1.2 nm (242 particles). The quaternary structure of AaLS-switch-ox was also examined by analytical ultracentrifugation. The sedimentation equilibrium behaviors of AaLS-switch-ox and wild-type AaLS are similar to each other and consistent with a 60-subunit stoichiometry (Figure 3.5). Fitting the data to a single ideal species model gave molecular weights of 920 ± 40 kDa and 970 ± 60 kDa for AaLS-switch-ox and wild-type AaLS, respectively, which are both close to the value calculated for a 60-subunit capsid (1010 kDa). Thus, thiophenol modification acts as a molecular switch for the conversion of pentamers into wild-type-like dodecahedral capsids.

3.2.3. Secondary Structure and Thermostability of AaLS-switch-ox

The structure and stability of the reassembled capsid were examined by circular dichroism (CD) spectroscopy (Figure 3.6). The far-UV CD spectra of pentameric AaLS-switch-red and AaLS-switch–ox capsid were nearly identical, indicating that capsid assembly does not involve changes in secondary structure. Further, wild-type AaLS gave a spectrum with a similar shape to that of the variants, which suggests a similar overall fold, but with a somewhat higher signal intensity (<20%), which may reflect a slightly

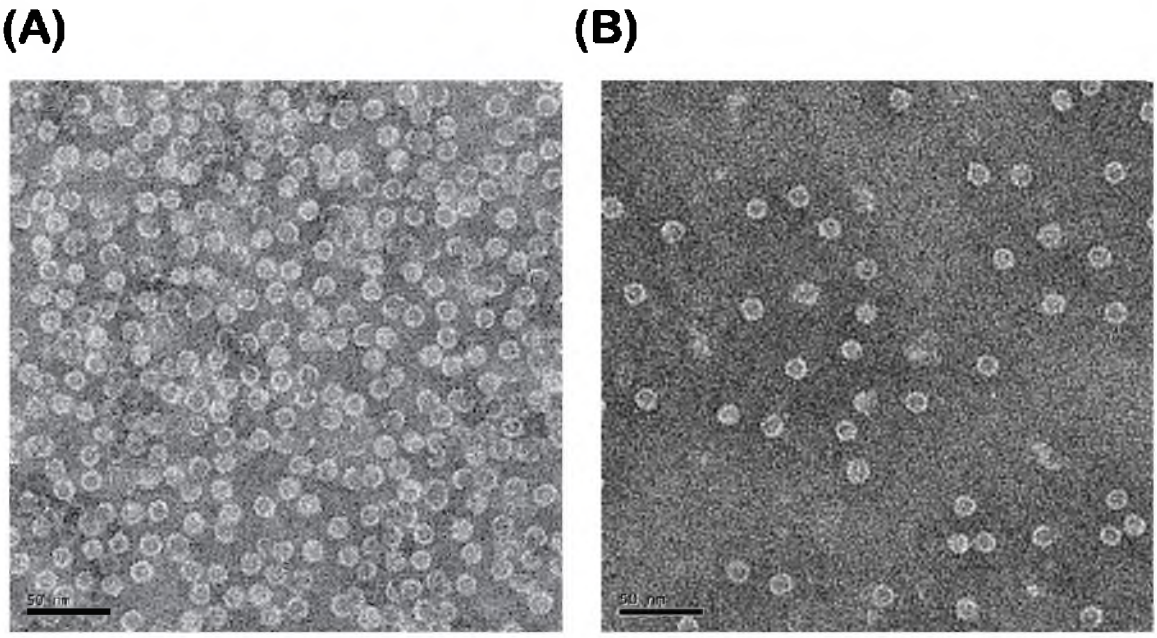


Figure 3.4. Visualization of AaLS capsids. Negatively stained transmission electron microscopy images of wild-type AaLS (A) and AaLS-switch-ox (B). Samples were treated with a staining solution containing 2% phosphotungstic acid at pH 8.0. 50-nm scale bars are shown as black rectangles.

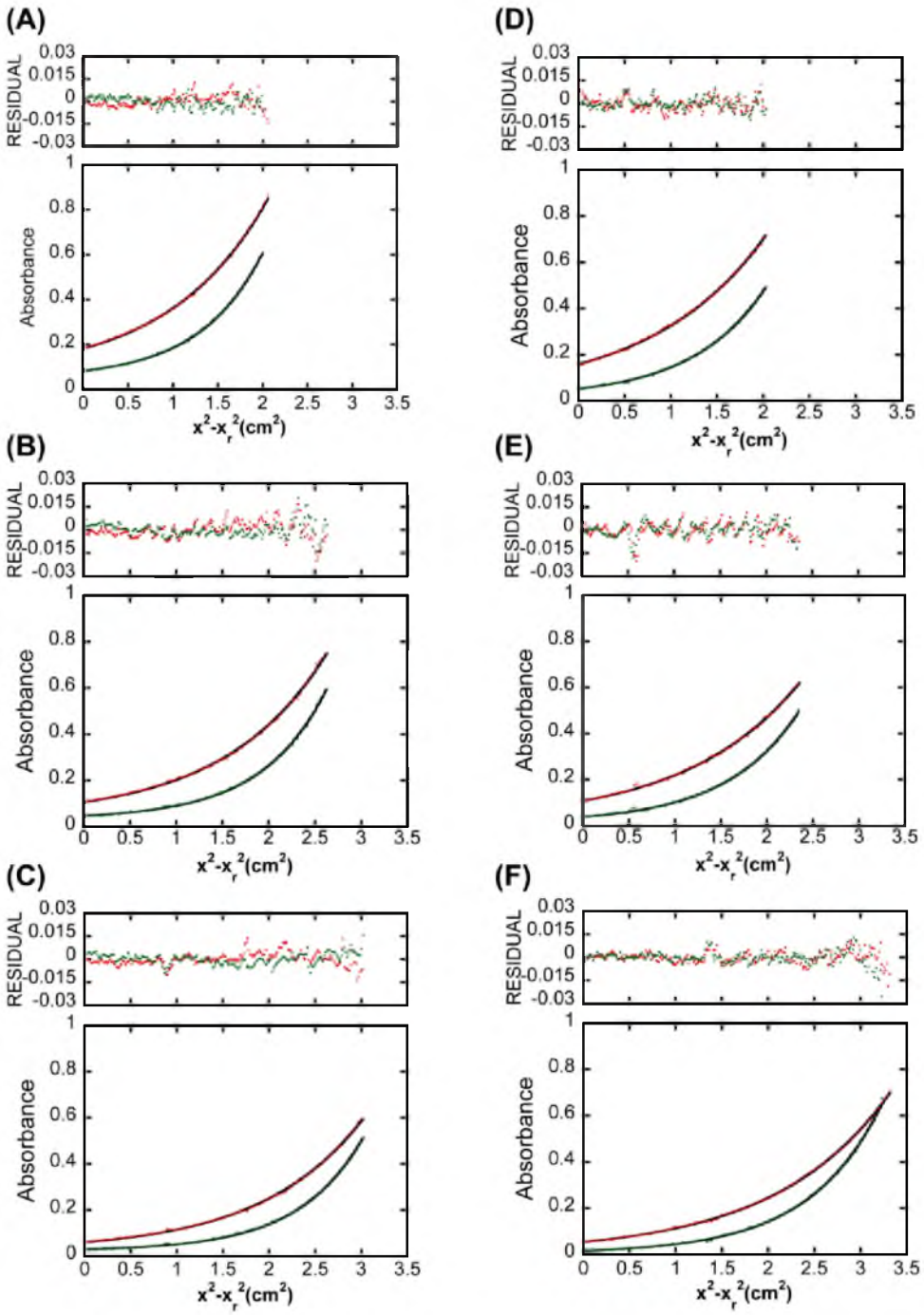


Figure 3.5. Plots of sedimentation equilibrium data. All samples were spun at two different speeds: 4000 rpm (red curves) and 5000 rpm (green curves). Plots A through C correspond to wild-type AaLS at 0.5 mg/mL (A), 0.25 mg/mL (B), and 0.13 mg/mL (C). Plots D through F correspond to AaLS-switch-ox capsids at 0.5 mg/mL (D), 0.25 mg/mL (E), and 0.13 mg/mL (F). Ideal curves resulting from fitting the data to a single ideal species model are shown in black.

Figure 3.6. Protein secondary structure and stability analysis by CD spectroscopy. (A) Far-UV spectra for wild-type AaLS (red), AaLS-switch-red (green), and AaLS-switch-ox capsids (blue). The spectra were recorded at 20 °C. (B) Melting curves for wild-type AaLS (red) and AaLS-switch-ox capsids (blue). The protein concentrations were 5 μM (monomer basis), and the cuvette pathlength was 1 cm in all cases.

lower secondary structure content in the AaLS-switch variants. The thermal melting curve of AaLS-switch-ox shows that the reassembled capsid is thermostable as no unfolding transition is apparent below 87 °C. However, the reassembled capsid has lost some stability relative to the wild-type AaLS capsid, which shows no change up to 95 °C (consistent with its reported melting temperature of 120 °C).[103] Nonetheless, AaLS-switch-ox retains a great deal of the structure and stability present in the wild-type capsid.

3.2.4. Timecourse Study of AaLS-switch-ox Capsid Assembly

Knowledge of the timecourse for AaLS-switch-ox assembly could help further optimize the capsid production method. Therefore, we measured the yield of capsid formation by performing SEC at various times after the addition of PEG-3350 to the protein (Figure 3.7). After 1 hour, the major peak in the chromatogram corresponds to the assembled 60-subunit AaLS-switch-ox capsid. Longer assembly times, up to 48 hours, gave even more capsid, up to ~87% yield. Between 48 hours and 96 hours, the yield of capsid did not significantly increase suggesting that the optimum assembly time is 48 hours.

Interestingly, the marginal increase in capsid yield seems to drop sharply after one hour (Figure 3.7). For example, the percent capsid yield increases from 60% to 72% between the 1 hour and the 4 hour time points, but the capsid yield only increases from 72% to 83% after another 20 hours of assembly time. The dramatic decrease in the rate of capsid formation at the later time points likely reflects a steep concentration dependence for this self-assembly process. This type of behavior is consistent with what has been observed for *in vitro* virus capsid assembly.[38]

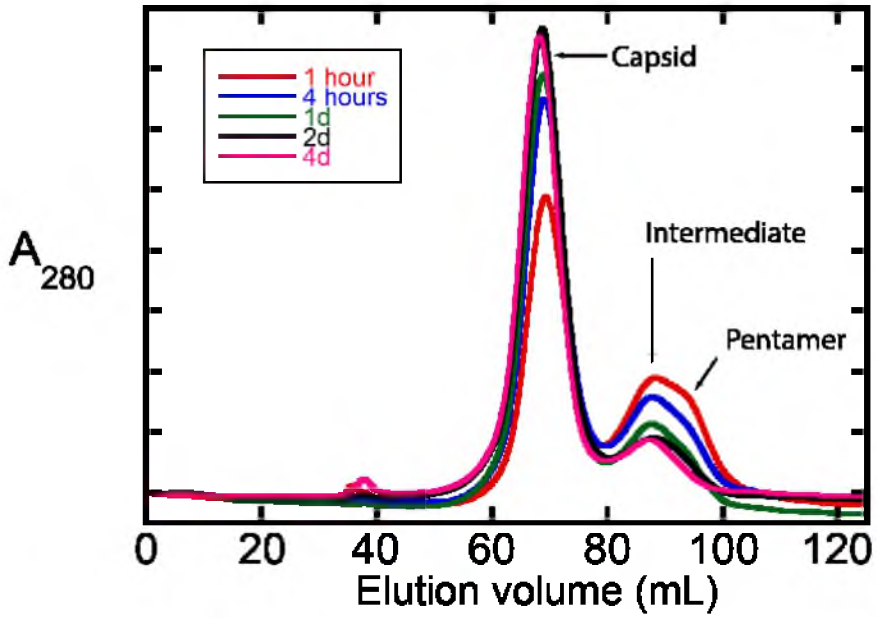


Figure 3.7. Timecourse study of AaLS-switch-ox capsid assembly. Size-exclusion chromatograms for protein samples injected at 1 h (red), 4 h (blue), 24 h (green), 48 h (black), and 96 h (pink) after the addition of PEG-3350. The positions of the peaks for 60-subunit capsid, assembly intermediate, and pentamer are indicated by arrows. The total protein concentration was 0.6 mg/mL. Assembly reactions were incubated at room temperature and analyzed by SEC at 4 °C.

An intermediate was also observed during the AaLS-switch-ox assembly timecourse. At the earliest time point (1 h), the pentameric starting material no longer gives a distinct peak, but the pentamer is still apparent as a shoulder on a new, earlier peak (Figure 3.7). This new peak presumably corresponds to an intermediate assembly state. At later time points, the shoulder disappears completely; the two remaining peaks represent this intermediate and the capsid. A more detailed understanding of the capsid assembly mechanism will require further study.

3.2.5. Structural Requirements for the Formation of the Three-fold Symmetric Interface in the Capsid

The thiophenol-modified cysteine of AaLS-switch-ox occurs at position 125, which is an isoleucine in wild-type AaLS. While the overall structure of AaLS-switch-ox and wild-type AaLS capsids are quite similar, thiophenol-modified cysteine is not very isosteric with isoleucine. We therefore examined the structural requirements of this modification for capsid assembly (Figure 3.8). The formation of disulfide adducts between the cysteine and ethanethiol, 1-propanethiol, or 2-propanethiol gives structures that more closely resemble isoleucine. However, none of these reagents were able to induce 60-subunit capsid formation upon modification of AaLS-switch-red and subsequent treatment with PEG-3350. Rather these adducts gave a mixture of pentamers and assemblies that were intermediate between pentamers and 60-subunit capsids. While it is not completely clear why thiophenol modification produces capsid but these aliphatic thiol reagents do not, this difference may be attributable in part to the greater nonpolar surface area of thiophenol.

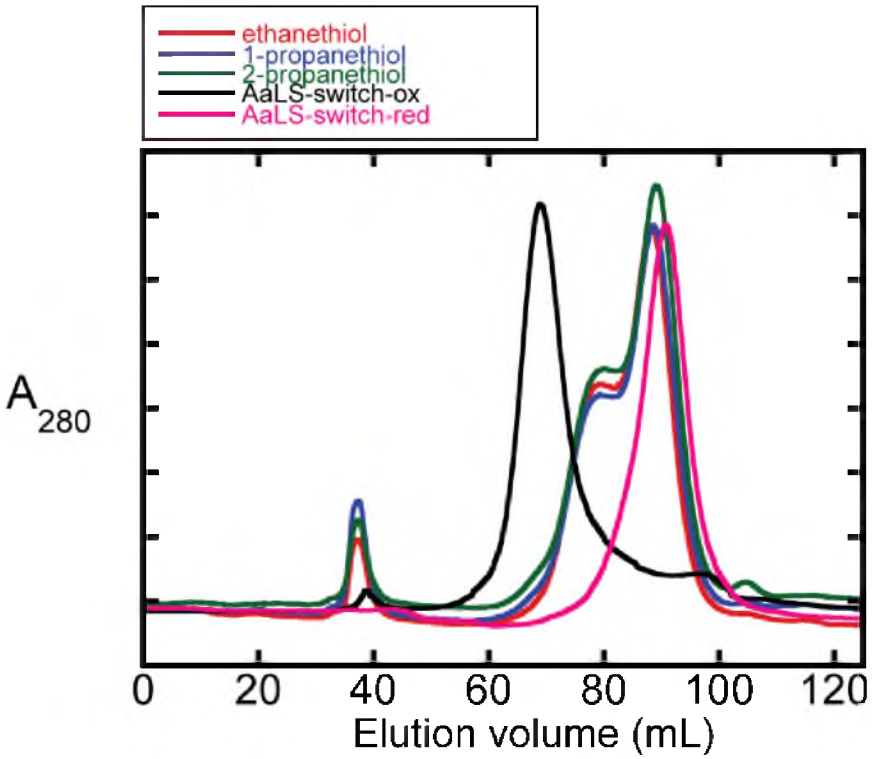


Figure 3.8. The influence of modification agent structure on assembly. AaLS-switch-red was treated with DTNB and then a variety of thiols using the procedure for modification of the protein with thiophenol described in the Materials and Methods. The resulting mixed disulfide adducts were treated with PEG-3350 according to the standard capsid assembly procedure for AaLS-switch-ox and then analyzed by SEC. Chromatograms are shown for the modification of the protein with ethanethiol (red), 1-propanethiol (blue), and 2-propanethiol (green). AaLS-switch-ox (black), which represents thiophenol modified protein, and unmodified AaLS-switch-red (pink) are shown as standards for comparison.

To further investigate the possible interactions of the thiophenol groups in the reassembled AaLS-switch-ox capsid, we performed molecular modeling. Starting with the crystal structure of wild-type AaLS (PDB id: 1HQK),[103] Ile125 was replaced with thiophenol-modified cysteine *in silico* using SYBYL-X 2.0. An energy minimization was carried out with the AMBER7 FF02 force field. The resulting model is free of steric clashes and shows only small deviations among the unchanged side-chains (0 to 1.2 Å) relative to the wild-type AaLS crystal structure (Figure 3.9). Thus, it seems possible to accommodate the thiophenol groups in the three-fold symmetric pore without invoking large scale restructuring of the protein. Nonetheless, the thiophenol groups in the model are in closer contact with each other than the corresponding isoleucines in the wild-type AaLS crystal structure. For example, the closest interatomic distance between the isolecuines in wild-type AaLS is ~4.3 Å, while in the model, there are several C-C and C-S distances in the range of 3.2 to 3.5 Å between thiophenol groups. In addition, the phenyl ring of thiophenol makes new contacts to the side-chains of L26 and L121 (in adjacent pentamers) and T130 (in the same subunit). Interestingly, the disulfide bonds of the three modified cysteines in the model have dihedral angles in the range of 62° to 66°. As the optimum dihedral angle for a disulfide bond is around 90°,[151] the favorable packing of the thiophenol groups may come at the expense of some torsional strain. Even though AaLS-switch-ox forms similar capsids to wild-type AaLS, the details of side-chain interactions at the three-fold symmetry axis may differ significantly.

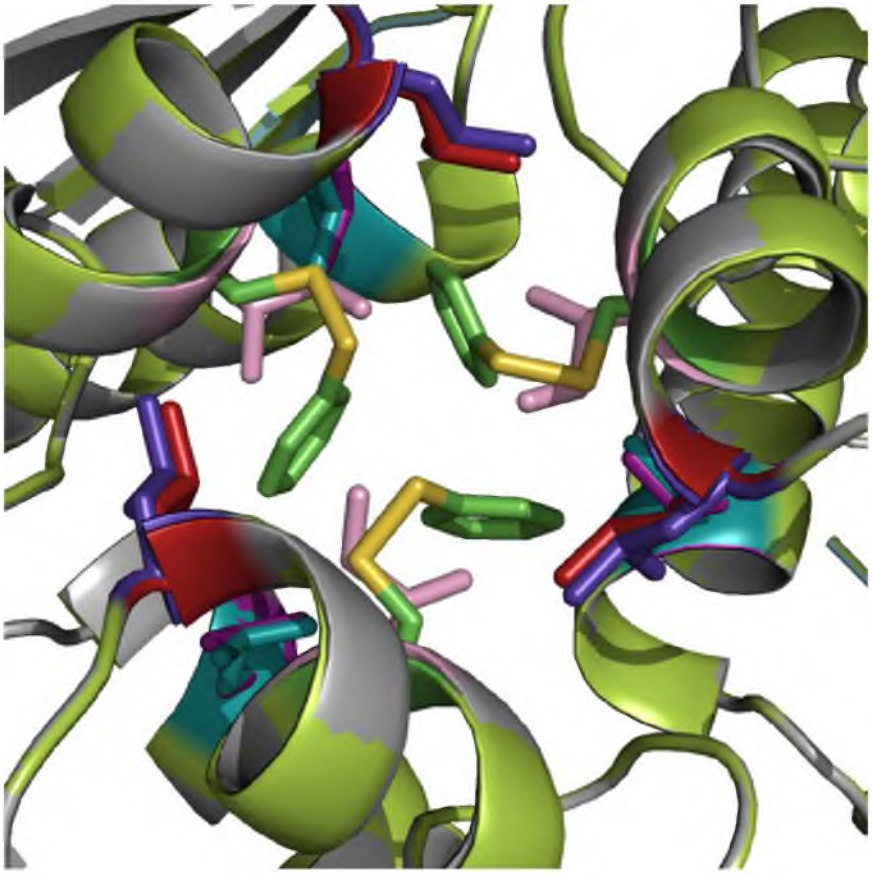


Figure 3.9. Overlay of the three-fold symmetric pores of wild-type AaLS and an energy minimized model of AaLS-switch-ox. Energy minimization was carried out for the AaLS-switch-ox (green ribbon) model using the program SYBYL-X 2.0 with the AMBER7 FF02 force field. The coordinates for the wild-type AaLS (grey ribbon) structure are from PDB id: 1HQK. In AaLS-switch-ox, the thiophenol-modified cysteine at position 125 is colored green and yellow, while L26 and L121 are colored cyan and purple, respectively. In the wild-type AaLS structure, I125 is colored pink, while L26 and L121 are colored magenta and red, respectively.

3.2.6. Attempts to Disassemble the AaLS-switch-ox Capsid by Disulfide Bond Reduction

The assembly of AaLS-switch-ox was triggered by covalent modification of C125 with thiophenol via disulfide bonding. In principle, this modification should be reversible upon treatment with reducing agents. Upon reduction, it is likely that the thiophenol would rapidly diffuse away from the protein causing the capsid to spontaneously dissociate into free pentamers. To test this hypothesis, AaLS-switch-ox capsid was incubated overnight in pH 8.0 buffer with various reducing agents (dithiothreitol (DTT), 2-mercaptoethanol (β-ME), tris (2-carboxyethyl) phosphine (TCEP), and $NaBH_4$). Surprisingly, the capsid was unaffected in all cases, based on SEC analysis (Figure 3.10A). The failure to disassemble was likely a consequence of the inability of the reducing agents to break the disulfide bond. ESI-MS analysis confirmed the lack of reduction by $NaBH_4$ (Figure 3.11). The major mass peaks of both untreated and $NaBH_4$-treated AaLS-switch-ox (16895.2 and 16896.8, respectively) were similar to that calculated for intact AaLS-switch-ox (16896). A small mass peak was observed at 16787.2 for the $NaBH_4$-treated sample, which is close to that expected for AaLS-switch-red (16787); however, this species might not represent a product of reaction with $NaBH_4$, as untreated AaLS-switch-ox also shows a small mass peak in this region (16784.8).

The lack of reaction may stem from burial of the disulfide within the protein, making it sterically inaccessible. A further attempt at reduction of AaLS-switch-ox capsid was carried out by incubation with DTT at 80 °C, with the hope that this high temperature would loosen the protein structure. Indeed, some pentamer was apparent after this treatment, but most of the protein remained in the capsid form (Figure 3.10B). The highly

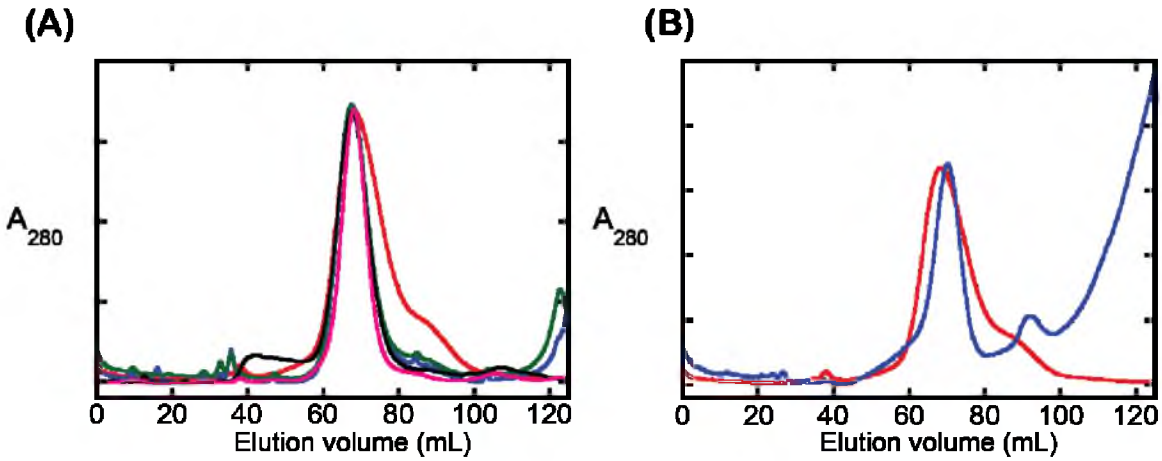


Figure 3.10. SEC analysis of AaLS-switch-ox capsids after treatment with reducing agents. (A) The AaLS-switch-ox capsid was treated with different reducing agents, and the assembly states were analyzed by SEC on a HiPrep 16/60 Sephacryl S-400 HR column at 4 °C in SEC running buffer (50 mM sodium phosphate, 200 mM NaCl, pH 8.0). The AaLS-switch-ox capsid was incubated with either 10 mM DTT (blue curve), 10 mM β-ME (green curve), 10 mM TCEP (black curve), or 10 mM $NaBH_4$ (magenta curve). Reactions were carried out by overnight incubation of protein and reducing agent in SEC running buffer at room temperature, except for the reaction with $NaBH_4$, which was performed at 37 °C. Untreated AaLS-switch-ox capsid (red curve) is shown as a standard for comparison. (B) AaLS-switch-ox capsid was incubated with 10 mM DTT in SEC running buffer at 80 °C for 60 min, and then the assembly state was analyzed by SEC on a HiPrep 16/60 Sephacryl S-400 HR column (blue curve). Untreated AaLS-switch-ox capsid is shown as a standard for comparison (red curve). Absorbance values were normalized for ease of visualization in all cases.

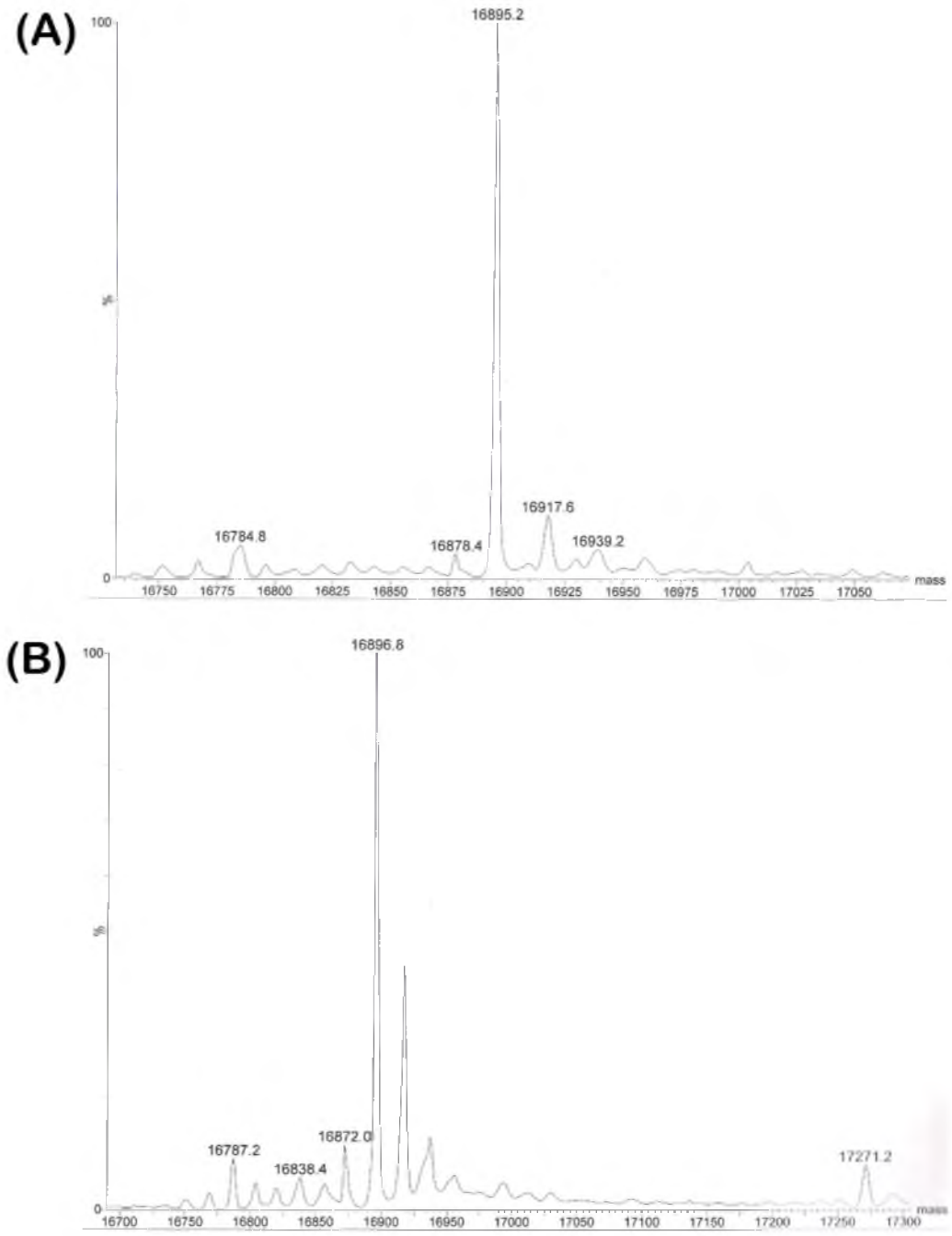


Figure 3.11. ESI-MS analysis of AaLS-switch-ox reduction by $NaBH_4$. (A) Mass spectrum of untreated AaLS-switch-ox. (B) Mass spectrum of AaLS-switch-ox following incubation with $NaBH_4$.

hydrophobic environment provided by the thiophenol group at the three-fold symmetry axis of the AaLS-switch-ox capsid may hinder access of the reducing agents to the disulfide bonds, making them extraordinarily unreactive. Regardless, more studies will be required to identify conditions that give efficient reduction and disassembly of the AaLS-switch-ox capsid.

3.3.Conclusions

We have exploited knowledge of pentamer-pentamer interactions in AaLS to build a capsid assembly switch. In AaLS-switch-red, site-directed mutagenesis was used to remove a nonpolar interaction at the three-fold symmetry axis and electrostatic interactions at the two-fold symmetry axis of the wild-type AaLS capsid halting assembly of the variant at the pentameric stage. Covalent modification of AaLS-switch-red with thiophenol adds nonpolar surface area at position 125, which drives the modified pentamers to associate into wild-type-like capsids. While this modification recapitulates an important feature of the wild-type AaLS capsid (hydrophobic clustering at the three-fold axis), the side-chain packing may differ considerably. Given that hydrophobic residues are commonly found to act as energetic hot spots for protein-protein interactions,[27] our cysteine modification strategy may provide a generally applicable method for controlling protein association.

Protein capsids can act as molecular containers that are useful for a variety of applications including catalysis, nanomaterials synthesis, bioimaging, and drug delivery.[60,67,76,118,147] The AaLS capsid shows great potential for these applications and has been used to encapsulate both proteins and RNA *in vivo*.[111–113] Guest loading *in vitro*

would potentially allow for the packaging of a wider variety of guest molecules (such as peptides/proteins, nucleic acids, synthetic polymers, nanoparticles, or small-molecules) as well as greater control over the stoichiometry and specificity of guest loading. However, wild-type AaLS lacks a controllable assembly switch, which presents a barrier to *in vitro* guest encapsulation. Nonetheless, a laboratory-evolved variant of AaLS was found to encapsulate highly cationic guest proteins *in vitro*.[114] The thiophenol-based capsid assembly switch presented here opens the door to the development of even more versatile *in vitro* encapsulation systems based on the AaLS scaffold. The ability to load a diverse set of guest molecules into protein capsids *in vitro* should provide a foundation for future applications in biotechnology, materials science, and nanomedicine.

3.4.Materials and Methods

3.4.1. Materials

Cell culture media and chemical reagents were obtained from the following suppliers: Sigma-Aldrich, Thermo Fisher scientific, Gold biotechnology, and MO-bio laboratories (unless otherwise noted). All purchased chemicals were used without further purification in this study. *Pfu*-turbo DNA polymerase and *E. coli* strains BL21 (DE3) and XL1-Blue were purchased from Stratagene. DpnI endonuclease was purchased from New England Biolabs. The oligonucleotides used in this chapter were synthesized by the DNA/peptide synthesis Core Facility at the University of Utah. UV-vis absorbance data were recorded using a Nanodrop 1000 spectrophotometer.

3.4.2. DNA Mutagenesis

AaLS-switch-red is a pentameric variant that contains a single cysteine per polypeptide at position 125. To generate AaLS-switch-red, a double site-directed mutagenesis PCR was performed using *Pfu*-turbo DNA polymerase according to the manufacturer's instructions and the detailed PCR parameters that were reported in Chapter 2. The mutagenic primers are listed in Table 3.2. The plasmid pMG-AaLSNoHis-R40S/H41S/I125S encodes a pentameric triple serine variant of AaLS (R40S/H41S/I125S-AaLS) and was used as the DNA template. The resulting plasmid, pMG-AaLSNoHis-switch-red, encodes C37A/R40S/H41S/I125C-AaLS and was confirmed by DNA sequencing (University of Utah DNA sequencing core facility).

3.4.3.Protein Production and Purification

The plasmid encoding the desired variant (pMG-AaLSNoHis or pMG-AaLSNoHis-switch-red) was transformed into $CaCl_2$-competent *E.coli* BL21 (DE3) cells. Protein overproduction was carried out using the procedure described in Chapter 2. The resulting cell pellets were stored at -20 °C until use.

Immediately before protein purification, the cells were resuspended in 10 mL of lysis buffer (50 mM sodium phosphate and 300 mM NaCl at pH 8.0) containing lysozyme (5mg), RNase A (0.6 mg), DNase I (10 μg), $MgCl_2$ (10mM), and DTT (10 mM). Following a 30-min incubation at room temperature, the cells were lysed by sonication. The cell lysate was then incubated at 60 °C for 5 min. The insoluble debris was removed by centrifugation (8000 x *g* for 30 min at 4 °C). The wild-type AaLS and AaLS-switch-red proteins were purified from the cleared supernatants using anion-exchange and size-

Table 3.2. List of mutagenic oligonucleotides used in site-directed mutagenesis PCR. Mutagenic bases are in bold.

Primer name	DNA sequence (5'→ 3')
HCN91	CCTTGGAACAGGCT**TG**CGAGCGCGCCGGCAC
HCN92	GTGCCGGCGCGCTCG**CA**AGCCTGTTCCAAGG
HCN105	GAGGGTGCAATTGAT**GC**CATAGTC**AGCAGC**GGCGGC
HCN106	GCCGCC**GCTGCT**GACTATG**GC**ATCAATTGCACCCTC

exclusion chromatography according to reported procedures. Typical yields were 32–40 mg protein per L of cell culture, as determined using the Coomassie Plus (Bradford) assay reagent (Thermo Scientific) according to the manufacturer's instructions.

3.4.4. Assembly of AaLS-switch-ox Capsids *In Vitro*

AaLS-switch-ox was produced from AaLS-switch-red by a two-step thiol-disulfide exchange process. Briefly, 1.0 mL of purified AaLS-switch-red protein (1.0–1.5 mg/mL in 50 mM sodium phosphate and 200 mM NaCl at pH 8.0) was gently dripped (~1 drop per s) into 1.0 mL of 20 mM 5,5′-dithiobis-(2-nitrobenzoate) (DTNB) dissolved in the same buffer (~200–300 equivalents of DTNB per protein thiol), which immediately turned bright yellow, indicating the splitting of the symmetric DTNB molecule into free 2-nitro-5-thiobenzoate (NTB) and a mixed disulfide adduct between NTB and C125 of the protein. After the addition of protein to DTNB was complete, the reaction was incubated for a further 15 min with stirring (~500 rpm) at room temperature. Measurement of both the absorbance at 412 nm (ε_{NTB} = 14,150 $M^{-1}cm^{-1}$)[152] and the protein concentration (as described above) gave a ratio of 1.1 mol NTB produced per mol of protein. The formation of the NTB-protein adduct was confirmed by electrospray mass spectrometry using a Waters LCT XE Premier ToF mass spectrometer (Table 3.1). To remove excess DTNB and NTB molecules, the reaction mixture was dialyzed at 4 °C for 12–16 h into 2 L buffer (50 mM sodium phosphate and 20 mM NaCl, pH 8.0). The protein solution was then moved to fresh dialysis buffer (2 L) and dialyzed for an additional 20–24 h.

The resulting NTB-protein conjugate was treated with 10 equivalents of thiophenol or

other thiols (0.4 M stock solution in DMF) with stirring (~500 rpm) at room temperature. The solution immediately turned bright yellow and was allowed to react for an additional 10–15 min. Measurement of both the [free NTB] (by $A_{412\,nm}$ as described above) and the [protein] (as described above) gave a ratio of 0.85 mol NTB released per mol of protein. The formation of the thiophenol-protein adduct was confirmed by electrospray mass spectrometry using a Waters LCT XE Premier ToF mass spectrometer (Table 3.1). The NTB-protein adduct was not found in the mass spectrum. Excess thiophenol and NTB molecules were removed via a two stage dialysis as described above. To trigger capsid assembly, the thiophenol-protein conjugate, AaLS-switch-ox, was spun in a 30 kDa MWCO Vivaspin concentrator (GE Healthcare) to a final volume of 1 mL and then transferred into a small glass vial containing 0.8 mL of PEG-3350 (25% w/v, dissolved in dialysis buffer). PEG-3350 has been shown to induce the formation of bacteriophage ø29 procapsids *in vitro* from a mixture of monomeric capsid proteins and dimeric scaffolding proteins.[150] This mixture was incubated at room temperature with slow stirring (200 rpm) for up to 4 days to let the capsid form. The protein was then exchanged into buffer lacking PEG-3350 by ultrafiltration with a 30 kDa MWCO Vivaspin concentrator.

3.4.5. Analysis of Protein Assembly State by Size-exclusion Chromatography

The assembly states of AaLS-switch-red and AaLS-switch-ox were analyzed by SEC using an ÄKTA FPLC system (GE Healthcare). The running conditions and FPLC parameters were based on those given in Chapter 2. Briefly, 1.0 mL of protein (0.3–

1.0 mg/mL) was injected onto a HiPrep 16/60 Sephacryl S-400 HR column that had been equilibrated with running buffer (50 mM sodium phosphate and 200 mM NaCl, pH 8.0). The flow rate was 0.5 mL/min, and the running temperature was 4 °C. Assembly states were determined based on comparison to wild-type AaLS (1010 kDa, 60-subunit capsid) and R40S/H41S/I125S-AaLS (84 kDa, pentamer). All further experiments involving AaLS-switch-ox capsids were performed using protein samples that had been purified using this size-exclusion column.

3.4.6. Sedimentation Equilibrium Analysis

AaLS-switch-ox capsids and wild-type AaLS were analyzed by Beckman XL-A analytical ultracentrifuge equipped with An60Ti rotor and photoelectric scanner using the protocol reported in Chapter 2. Both samples were spun at 4000 and 5000 rpm and at three different concentrations (130 μg/mL, 250 μg/mL, and 500 μg/mL). The data were analyzed using Ultrascan (University of Texas Health Science Center) and fit to a single ideal species model, as previously described in Chapter 2.

3.4.7. Transmission Electron Microscopy and Analysis of Average Particle Diameters

Wild-type AaLS and AaLS-switch-ox capsids (0.2 mg/mL each in SEC running buffer) were applied to copper grids (200 square mesh formvar coated). After waiting for 1 min to allow adhesion, excess protein solution was wicked away with sterile filter paper. The samples were then immediately stained with 2% phosphotungstic acid (pH 8.0) for 1 min. After staining, excess staining solution was removed with filter paper. TEM

imaging of the negatively stained samples were obtained by Dr. Michael Standing (Brigham Young University) using an FEI Technai G2 transmission electron microscope. The resulting images were analyzed using ImageJ software (National Institutes of Health). After creating a threshold image with a dark background, the particles on this threshold image were measured by limiting the analysis to a circularity of 0.2–1.0 and an area of 100–500 nm^2. A total of 307 and 242 particles were analyzed for wild-type AaLS and AaLS-switch-ox, respectively.

3.4.8. Circular Dichroism Spectroscopy

The CD spectroscopy was performed as described in Chapter 2. Both the far-UV spectra and thermal denaturation curves were obtained using 5 μM protein solutions in 1-cm pathlength cuvettes.

3.4.9. Modeling the Three-fold Symmetry Axis of the AaLS-switch-ox Capsid

A model of the thiophenol group interactions in the AaLS-switch-ox capsid was made starting from a fragment of the crystal structure of wild-type AaLS (PDB id: 1HQK) consisting of three adjacent pentamers. In the program SYBYL-X 2.0 (Tripos), a new amino acid was defined corresponding to the structure of a mixed disulfide between thiophenol and cysteine. The I125 residues at the three-fold axis were then replaced by this newly defined amino acid. An energy minimization was performed using the Powell method and the AMBER7 FF02 force field over the region within 8 Å of the thiophenol-modified cysteine. The default parameters were selected with the following exceptions:

Charges (AMBER), Dielectric Function (Distance), Termination (Gradient, 0.005 kcal/(mol·A)), and Maximum Iterations (1,000,000). The resulting energy minimized structure was visualized using Pymol (DeLano Scientific).

3.4.10. Treatment of AaLS-switch-ox Capsid with Reducing Agents

To attempt disassembly of the capsid, AaLS-switch-ox was treated with reducing agents. The standard reaction mixture contained AaLS-switch-ox (1 mg/mL) and either DTT, β-ME, TCEP, or $NaBH_4$ (10 mM each) in pH 8.0 buffer containing 50 mM sodium phosphate and 200 mM NaCl. For DTT, β-ME, and TCEP the mixtures were incubated at room temperature overnight. For $NaBH_4$, the mixture was incubated at 37 °C overnight. A second reaction with DTT was carried out at 80 °C for 60 min. The assembly state of the protein was then analyzed by SEC as described above in section 3.4.5.

CHAPTER 4

AN ENGINEERED pH-DEPENDENT SWITCH FOR CAPSID DISASSEMBLY AND REASSEMBLY

4.1. Introduction

Proteins represent a fascinating class of building blocks for constructing supramolecular assemblies.[28,29,153] Many structures, including cyclic oligomers, linear filaments, and polyhedral capsids, can be built by self-assembling monomeric protein subunits.[1] Polyhedral protein capsids are a particularly interesting type of supramolecular structure that forms a closed surface via the hierarchical self-assembly of lower order oligomers.[4,41]

The hollow interiors of capsids can be used for storage, transport, or activity regulation of cargo molecules. For example, the protein capsids of viruses act to prevent degradation of their nucleic acid genomes and to deliver them into a new host cell.[4] Consequently, virus capsids possess the ability to switch their assembly states between higher and lower order oligomers in a controlled fashion which facilitates cargo loading and unloading.[154] The assembly states of many virus capsids can also be switched *in vitro* by changing the environmental conditions, including ionic strength, metal ion concentration, or pH.[82,85,87] For example, Hepatitis B virus capsid protein forms homodimers at low ionic strength and 240-subunit icosahedral capsids at high ionic strength.[85] The 360-subunit icosahedral capsid of the murine polyomavirus self-assembles from pentameric building blocks upon adding Ca^{2+} or decreasing the pH to 6.0.[87] CCMV capsid protein forms homodimers at pH 8.0, and these dimers self-assemble into 180-subunit icosahedral capsids at pH 5.0.[82] Given their enormous biodiversity, viruses undoubtedly use many more capsid assembly switches, but assembly/disassembly mechanisms remain difficult to identify.

Inspired by viruses, protein capsids have become attractive scaffolds to develop novel encapsulation systems for use in medicine and biotechnology.[61,76] Indeed, some protein

capsids have been engineered to generate drug carriers or nanoreactors by specifically encapsulating drug molecules or enzymes, respectively, into their inner spaces.[74,83] Loading cargo molecules into protein capsids can be achieved by either *in vivo* or *in vitro* approaches.

In vivo loading requires coproduction of cargo molecules and capsid proteins in the same cell.[79,84,111] In this approach, capsid loading is presumably carried out simultaneously with capsid assembly in the cell. However, the encapsulated cargoes can be contaminated by other undesired cellular components due to the complexity and crowdedness of the intracellular environment. Further, the requirement for coproduction of capsid and cargo molecules within the same cell restricts the available set of potential cargo molecules that is compatible with the *in vivo* loading approach.

On the other hand, cargoes can be loaded into a purified intact capsid *in vitro* by diffusion of cargo molecules across the capsid shell.[66] Because *in vitro* loading is performed under more controlled conditions compared to the intracellular environment, it usually permits cleaner cargo encapsulation, more flexibility in cargo selection, and better control over the ratio of cargo to capsid. When loading cargoes into an intact capsid, the cargo molecules need to pass through pores in the capsid shell in order to diffuse into the capsid interior. In some cases, cargoes can be too big to cross the capsid shell, which narrows the size-range of potential cargo molecules that can be loaded by this approach relative to *in vivo* loading.

The size limit for *in vitro* cargo loading can potentially be overcome when a capsid assembly switch is available. In this case, the capsid can assemble around its cargo, which allows for larger guest molecules compared to loading of intact capsids.[80,114] However,

assembly switches are not available for all capsids. Further, when a natural switch is available, the solution conditions that trigger capsid assembly are not always compatible with the desired applications. Therefore, it can be desirable to engineer assembly switches into capsids.

The protein capsid formed by AaLS (Figure 2.1) is a promising scaffold for developing novel encapsulation systems. Engineered AaLS variants possessing highly charged interiors can exploit charge complementarity to encapsulate either RNA or other proteins upon coproduction of capsid and cargo in *E. coli* cells.[111–113] Further, a negatively charged AaLS variant, AaLS-13, is also able to form capsids around high positively charged cargo molecules *in vitro*.[114] While AaLS capsids have demonstrated promising results in the development of novel encapsulation systems, the lack of a general method to control the self-assembly of AaLS capsids could limit the versatility of AaLS capsid for further applications.

In Chapter 3, a switch for AaLS capsid formation based on redox chemistry was developed. This pentameric AaLS variant, AaLS-switch-red, contains four mutations (C37A, R40S, H41S, and I125C), in which R40S, H41S, and I125C disrupt assembly, trapping this variant at the pentamer stage. The C37A and I125C mutations install a unique cysteine at a position corresponding to the three-fold symmetry axis of the wild-type capsid. Modification of the unique cysteine by disulfide bonding with thiophenol generates AaLS-switch-ox, triggering capsid formation by recapitulating a hydrophobic patch found in wild-type AaLS. However, the AaLS-switch-ox capsid is not readily disassembled.

In this chapter, a pH-dependent capsid assembly switch is reported. This assembly

switch was designed using knowledge of important interactions at the pentamer interfaces in the AaLS capsid (Figure 2.1). Three histidine residues per AaLS subunit were engineered near the hydrophobic patch at the three-fold symmetry interface. The titratable imidazole group in the side-chains of these histidines minimally disrupts the capsid assembly at pH 8.0. However, upon lowering the pH, these engineered histidines disrupt the hydrophobic interaction via charge repulsion and trigger capsid disassembly. After bringing the pH back to 8.0, this engineered AaLS protein reassembled into wild-type like capsids. This switch allows for reversible capsid dissociation and reconstruction, providing full control over assembly state.

4.2. Results

4.2.1. Design of AaLS-switch-pH

The important role of the interactions at the three-fold symmetric axis in AaLS capsid assembly has been demonstrated by the mutagenesis studies and the thiophenol modification studies described in Chapter 2 and Chapter 3, respectively. Therefore, it should be possible to alter the assembly state of AaLS by the strategic introduction of charge repulsion near the three-fold symmetric axis. To test this hypothesis, a variant of AaLS, named AaLS-switch-pH, was designed (Figure 4.1). In AaLS-switch-pH, three amino acids (T120, E122, and Q123) located around the three-fold axis at the interior surface of the capsid were simultaneously substituted by histidines. The imidazole side-chain of histidine can switch between neutral and positively charged states by pH titration. Thus, these mutations should destabilize the pentamer interface of AaLS-switch-pH at pH values below 6.0, where histidines are mostly charged, but minimally destabilize the

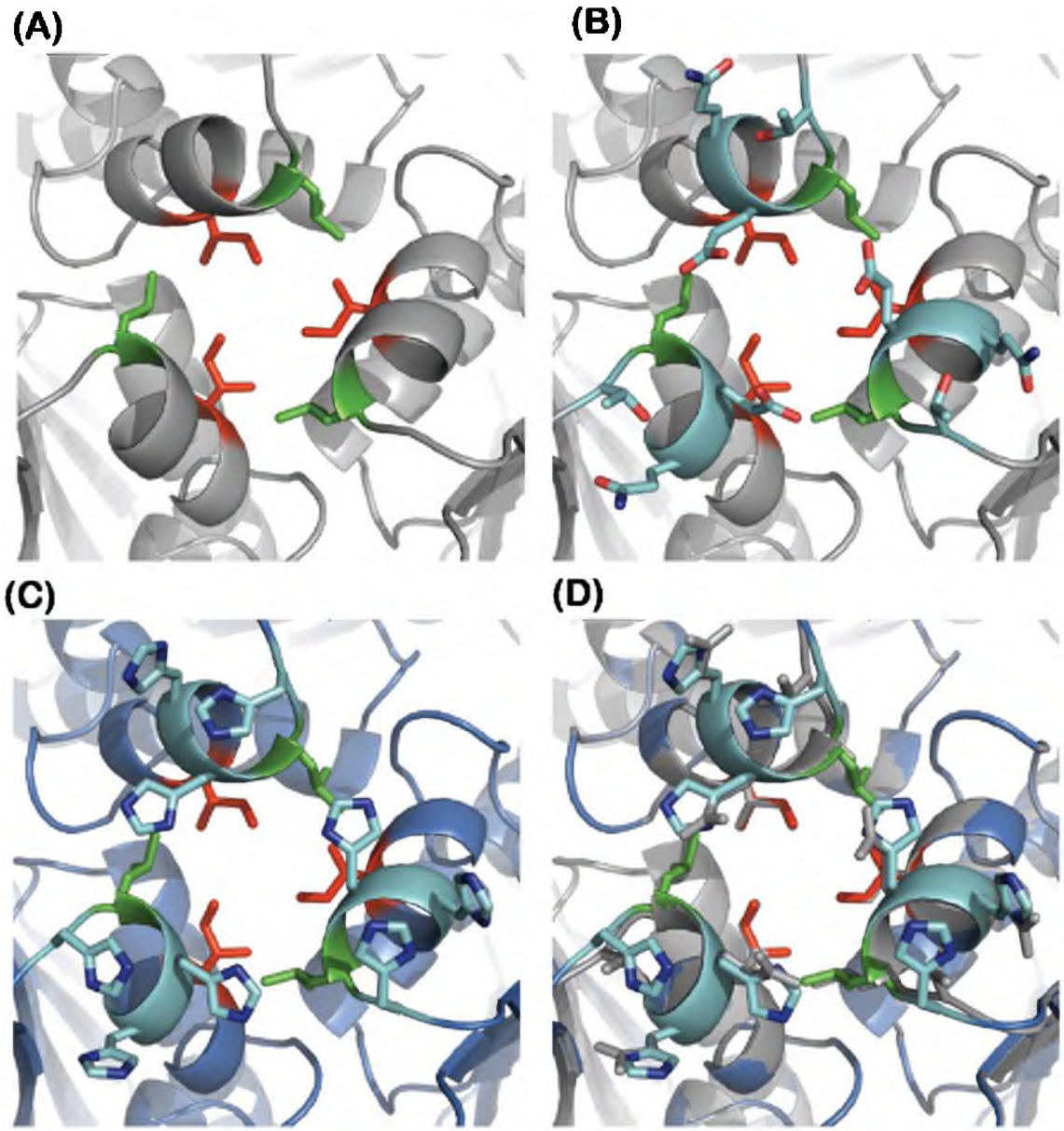
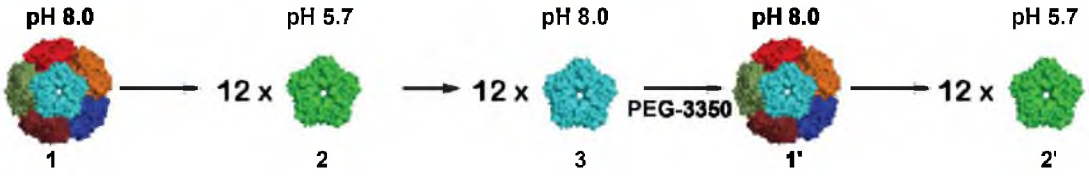


Figure 4.1. Summary of the AaLS-switch-pH design. Panels (A) to (D) are all viewed down the three-fold symmetry axis. (A) The hydrophobic interaction site at the three-fold symmetry axis of wild-type AaLS consists of L121 and I125, which are shown as sticks and colored in green and red, respectively. (B) Amino acids T120, E122, and Q123 neighbor the three-fold symmetry axis in wild-type AaLS and are shown as cyan-colored sticks. (C) An energy-minimized model of AaLS-switch-pH shows the newly introduced histidines at positions 120, 122, and 123 as sticks, which are colored cyan. (D) The energy minimized model of AaLS-switch-pH (blue ribbon, residue coloring as in panel C) is overlayed with wild-type AaLS (grey ribbon and residues T120, L121, E122, Q123, and I125 shown as grey sticks).

capsid at pH values above 7.0, where histidines are neutral. In addition to the three histidine mutations, R40S and H41S mutations were also incorporated into AaLS-switch-pH to weaken the pentameric interface at the two-fold symmetry axis.

4.2.2. AaLS-switch-pH Shows pH-dependent Disassembly and Reassembly *In Vitro*

AaLS-switch-pH protein was produced in *E. coli* cells and purified by the method reported in Chapter 3. All the purification steps were carried out in phosphate buffer at a constant pH of 8.0. SEC analysis shows similar elution behavior of freshly purified AaLS-switch-pH (**1** in Scheme 4.1) and wild-type AaLS suggesting that these two proteins form similar capsids during production *in vivo* (Figure 4.2 blue and red chromatograms). The capsid **1** was then dialyzed into citrate buffer at pH 5.7, giving **2** (Scheme 4.1). SEC analysis (Figure 4.2 green and blue chromatograms) showed that **2** (at pH 5.7) eluted much later than **1** (at pH 8.0), indicating that the capsid had disassembled. The elution behavior of **2** was similar to pentameric R40S/H41S/I125S-AaLS at pH 5.7 (Figure 4.2 orange chromatogram), which suggests that **2** is a pentamer. Interestingly, the AaLS-switch-pH capsid can be reassembled (**1′** in Scheme 4.1) by subsequent dialysis of **2** back into phosphate buffer at pH 8.0 and the addition of PEG-3350 (Figure 4.2 black chromatogram). Further dialysis of **1′** back into citrate buffer at pH 5.7 gave **2′**, which eluted similarly to **2** during SEC (Figure 4.2 magenta chromatogram), showing that this assembly/disassembly process is reversible. As a control, wild-type AaLS capsid was also subjected to the treatment with citrate buffer at pH 5.7. However, no significant change in the SEC elution behavior was observed (Figure 4.3), demonstrating that the pH-



Scheme 4.1. Disassembly and reassembly scheme for AaLS-switch-pH capsid. Species **1** corresponds to AaLS-switch-pH capsid that has been freshly purified from *E. coli* cells, whereas species **1′** corresponds to AaLS-switch-pH capsid that has been reassembled *in vitro*. Species **2** and **2′** correspond to the pentamers resulting from acidic disassembly of **1** and **1′**, respectively. Species **3** corresponds to pentameric AaLS-switch-pH at pH 8.0

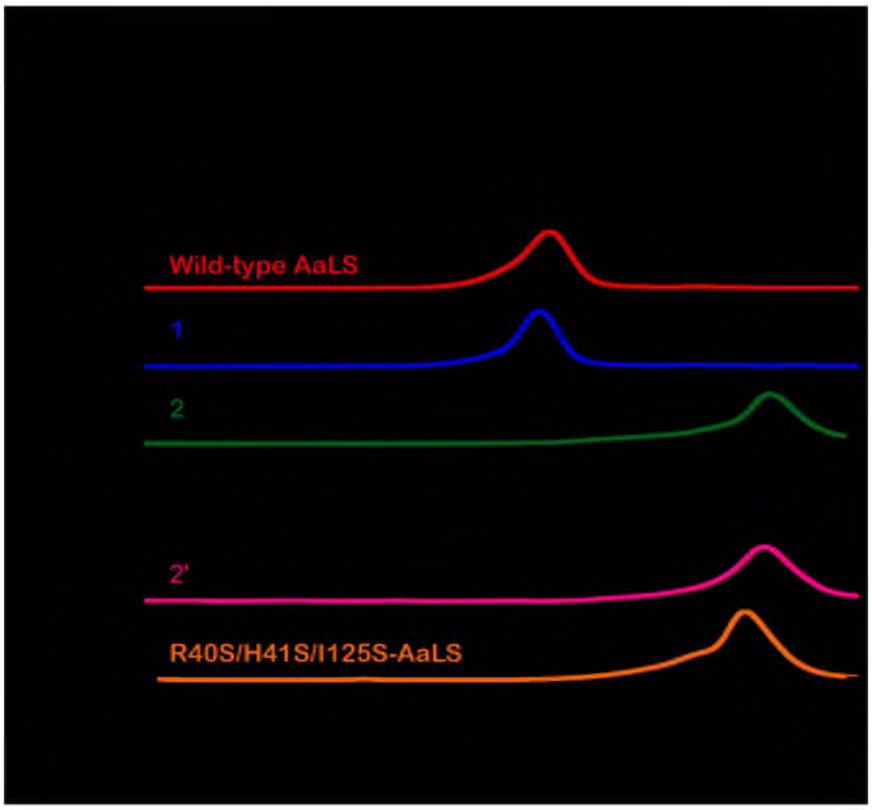


Figure 4.2. The reversible pH-dependent assembly/disassembly of AaLS-switch-pH assessed by SEC. Chromatograms are shown for wild-type AaLS (red), **1** (blue), **2** (green), **1′** (black), **2′** (pink), and AaLS-R40S/H41S/I125S (orange), which were each run over a HiPrep 16/60 Sephacryl S-400 HR column. A running buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0 was used for wild-type AaLS, **1** and **1′**. A running buffer containing 50 mM sodium citrate and 200 mM NaCl at pH 5.7 was used for **2**, **2′**, and R40S/H41S/I125S-AaLS. The absorbance values for all chromatograms were normalized to facilitate comparison.

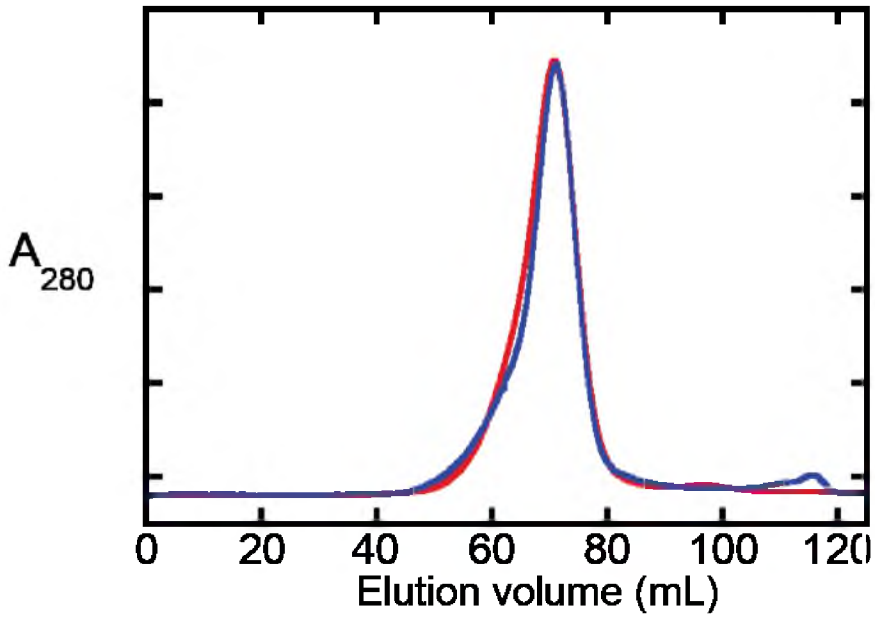


Figure 4.3. pH-independence of wild-type AaLS assembly. SEC was performed on samples of wild-type AaLS at different pH and buffer conditions using a HiPrep 16/60 Sephacryl S-400 HR column. The running buffer for the red trace contained 50 mM sodium phosphate and 200 mM NaCl at pH 8.0. The running buffer for the blue trace contained 50 mM sodium citrate and 200 mM NaCl at pH 5.7. The absorbance values were normalized for ease of comparison. The red trace shown above is identical to the red trace from Figure 4.2.

dependent shift in the assembly state of AaLS-switch-pH must be a consequence of the engineered histidine mutations.

The capsid structures of wild-type AaLS, **1**, and **1′** were visualized by negative stain transmission electron microscopy (TEM) (Figure 4.4 A–C). The TEM images showed that **1** and **1′** both formed similar spherical particles with average diameters of 16 ± 3 nm (measured from 127 particles) and 17 ± 4 nm (measured from 154 particles), respectively. Further, both **1** and **1′** capsids are similar in size and shape to the wild-type AaLS capsid (Figure 4.4 D–F). Visual inspection of the averaged particle images for wild-type AaLS, **1**, and **1′** suggests that **1′** has a lighter interior than the other two capsids. To confirm this observation, the gray value was plotted against the distance along the cross section of each capsid (Figure 4.4 G). While all three plots have roughly the same shape (with two peaks corresponding to the capsid walls and a local minimum in between indicating empty space inside the capsid where stain molecules have accumulated), **1′** is indeed lighter in the middle. Interestingly, the capsid **1** is even darker inside than both **1′** and wild-type AaLS. The different extents of staining the capsid interiors in these three samples could reflect how empty the capsids are. The presence of other molecules inside the capsid could leave less space for stain molecules to accumulate, giving less dark signal. In this case, capsid **1** would be emptier and capsid **1′** would be less empty than wild-type AaLS. It is not clear why **1** would be hollower. **1′** could possibly be partially filled with PEG-3350. Imaging artifacts may provide an alternative explanation for the variations in the darkness of the capsid interiors that would call into question the significance of these observed differences. Two factors that could possibly contribute to such artifacts are 1) differences in staining efficiency (both between and within EM grids)

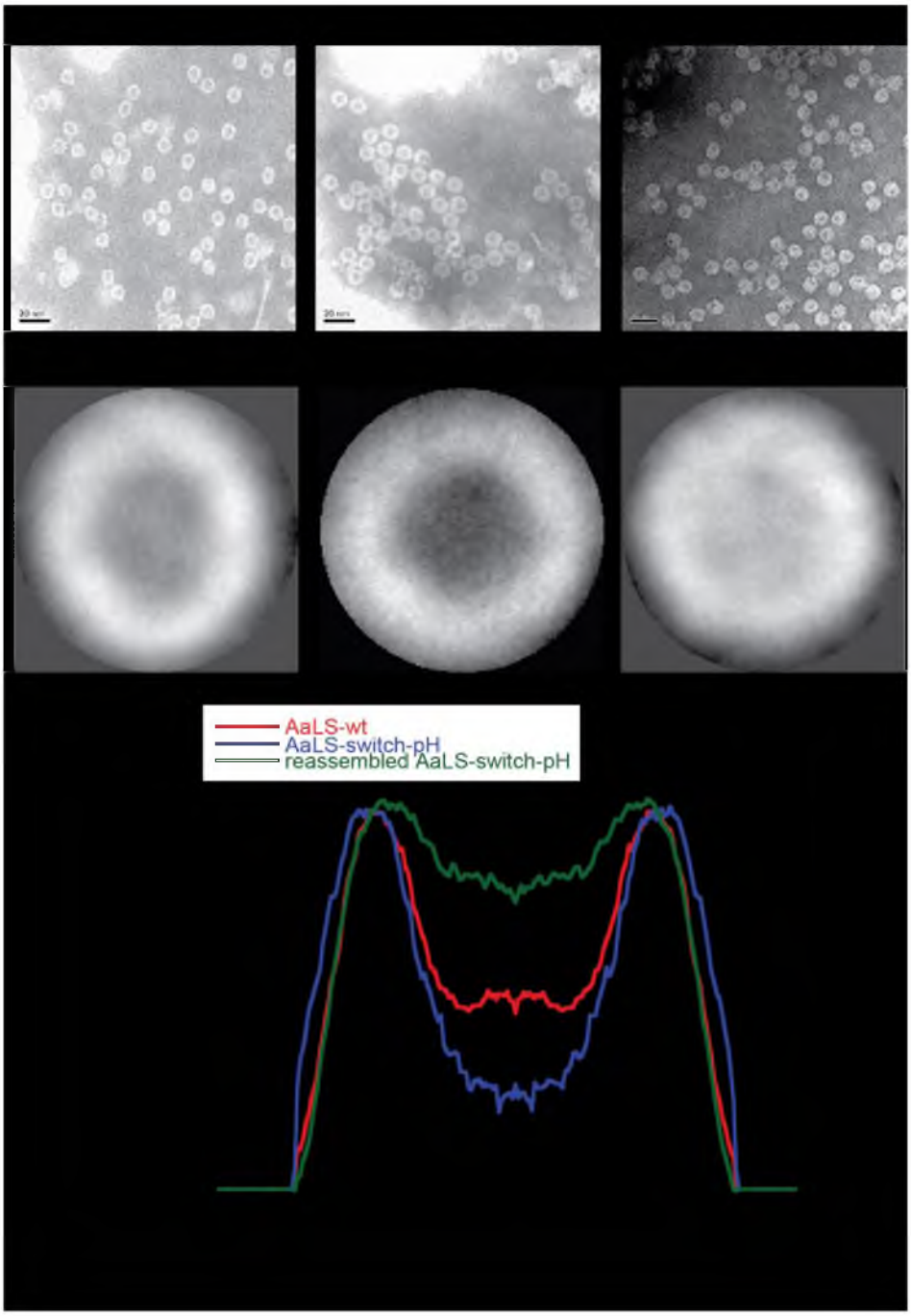


Figure 4.4. Visualization of AaLS capsids by electron microscopy. TEM images are shown for wild-type AaLS (A), **1** (B), and **1′** (C), which were each negatively stained with phosphotungstic acid. 20 nm scale bars are shown as black rectangles. Averaged TEM images of 46 particles of wild-type AaLS (D), 31 particles of **1**, (E), and 35 particles of **1′** (F). The images shown in panels D–F are all at the same scale. (G) Gray value versus distance across averaged capsid images. A higher gray value represent lighter area in the averaged images and a lower gray value represents darker area.

and 2) the relatively small sample sizes used in the averaging. Further studies will be required to resolve this issue. Nonetheless, both **1** and **1′** share a great deal of structural similarity both with each other and with wild-type AaLS.

The quaternary structures of AaLS-switch-pH at different pH values were also confirmed by sedimentation equilibrium analysis (Figure 4.5 and Table 4.1). The data obtained for **1** and **1′** at pH 8.0 fit well to a single ideal species model and gave molecular weights of 890 ± 60 kDa and 1000 ± 30 kDa, respectively. These molecular weights are close to the calculated value, 1013 kDa, for a 60-subunit assembly of AaLS-switch-pH based on its amino acid sequence. On the other hand, fitting the data obtained for **2** at pH 5.7 to the same model gave a molecular weight of 81 ± 1 kDa, which is close to the calculated molecular weight of a pentamer (84 kDa). The results from sedimentation equilibrium studies further support the notion that AaLS-switch-pH undergoes a reversible pH-dependent change in quaternary structure between a 60-subunit capsid at pH 8.0 and pentamers at pH 5.7.

4.2.3. Timecourse Study of AaLS-switch-pH Assembly

The studies of AaLS-switch-ox assembly (Chapter 3) indicate that the addition of PEG-3350 is crucial for efficient formation of dodecahedral capsids from pentamers *in vitro*. In this respect, AaLS-switch-pH behaves similarly. At a protein concentration of 0.5 mg/mL, overnight dialysis of the pentamer **2** into phosphate buffer at pH 8.0 gives **3** (Scheme 4.1), which shows no change in the assembly state (Figure 4.2 green chromatogram versus Figure 4.6A blue chromatogram). Further, only a small yield of capsids, **1′**, was obtained when the protein concentration was increased to 1 mg/mL and

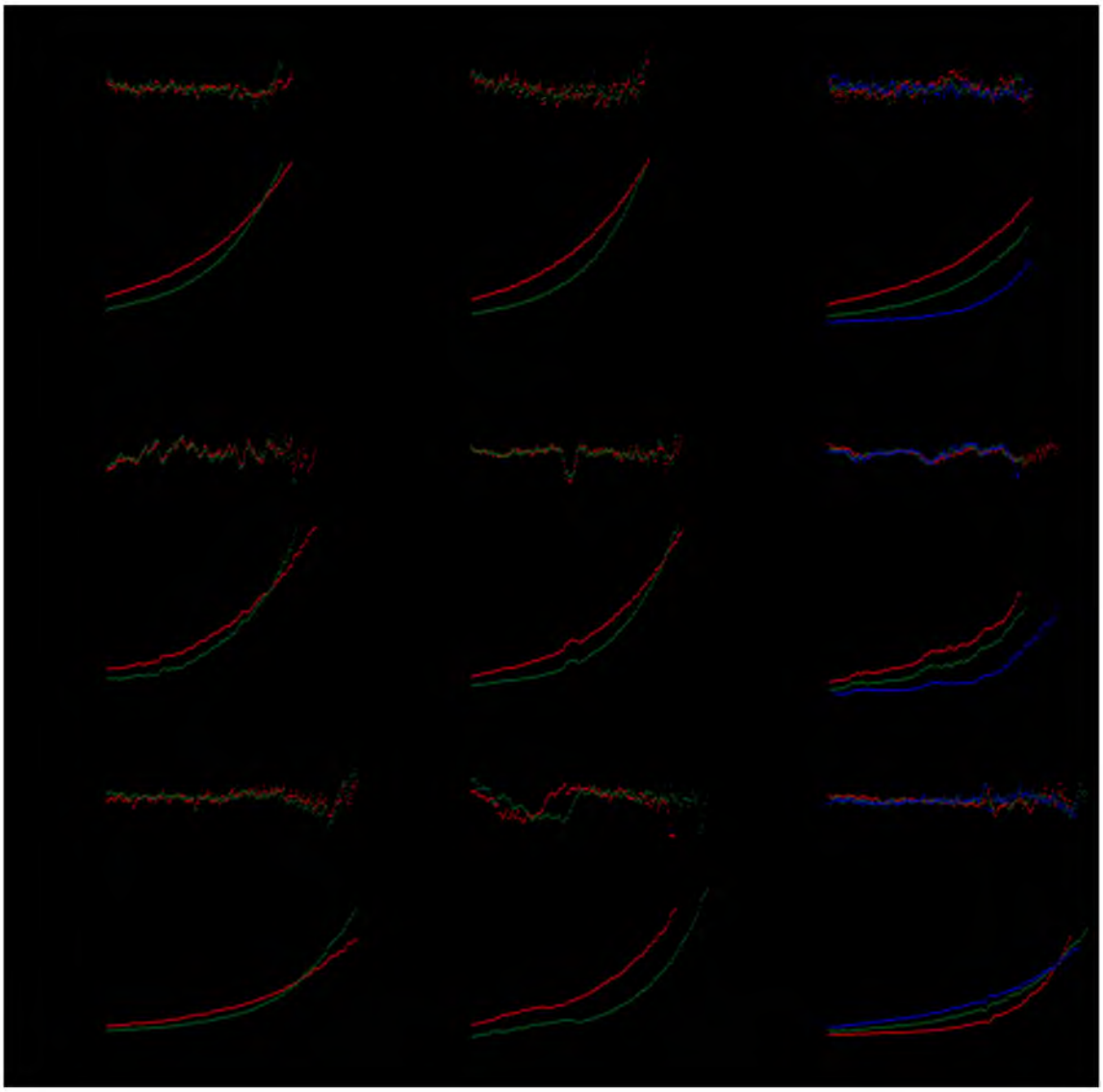


Figure 4.5. Plots of sedimentation equilibrium data. Plots A through C correspond to samples of **1** that had been dialyzed into buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0. These samples were spun at 4000 rpm (red curves) and 5000 rpm (green curves) at concentrations of 1.0 mg/mL (A), 0.75 mg/mL (B), and 0.5 mg/mL (C). Plots D through F correspond to samples of **1′** that had been dialyzed into buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0. These samples were spun at 4000 rpm (red curves) and 5000 rpm (green curves) at concentrations of 1.0 mg/mL (D), 0.75 mg/mL (E), and 0.5 mg/mL (F). Plots G through I correspond to samples of **2** that had been dialyzed into buffer containing 50 mM sodium citrate and 200 mM NaCl at pH 5.7. These samples were spun at 12500 rpm (red curves), 15000 rpm (green curves), and 20000 rpm (blue curves) at concentrations of 1.0 mg/mL (G), 0.75 mg/mL (H), and 0.5 mg/mL (I). In all cases, ideal curves resulting from fitting the data to a single ideal species model are shown in black.

Table 4.1. Assembly states of AaLS variants determined by sedimentation equilibrium analysis. Based on the amino acid sequence of AaLS-switch-pH, the calculated molecular weight for a 60-subunit capsid and a pentamer is 1013 kDa and 84 kDa, respectively.

Protein	Observed molecular weight (kDa)
AaLS-switch-pH (**1**)	890 ± 60
Disassembled AaLS-switch (**2**)	81 ± 1
Reassembled AaLS-switch-pH (**1′**)	1000 ± 30

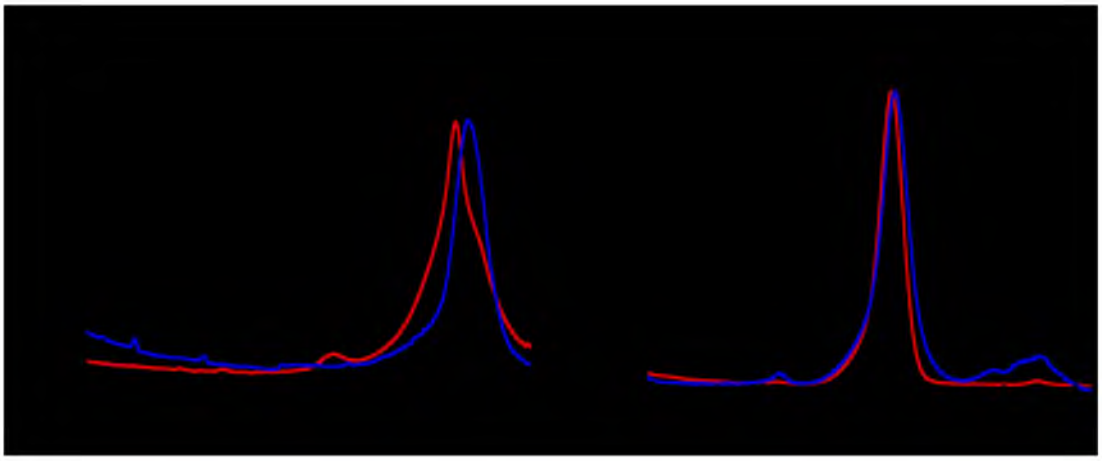


Figure 4.6. Long term stability of **3** and **1′** in the absence of PEG-3350. (A) Immediately following dialysis into buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0, a sample of **3** (blue) was injected onto a HiPrep 16/60 Sephacryl S-400 HR column that had been equilibrated with the same buffer. The major peak fractions were pooled, concentrated to 1 mg/mL, incubated for 4 weeks at 4 °C in the absence of PEG-3350, and then re-analyzed by SEC (red). (B) Immediately following assembly in the presence of PEG-3350, a sample of **1′** was purified by SEC, the capsid-containing fractions were pooled and concentrated to 1 mg/mL, and the protein was incubated for 3 weeks at 4 °C in the absence of PEG-3350. This 3-week-old sample of **1′** (red) was then injected onto a HiPrep 16/60 Sephacryl S-400 HR column. For comparison, a chromatogram of freshly purified **1′** is shown in blue.

the incubation time was prolonged to 4 weeks at 4 °C, although most of the protein eluted somewhat earlier than **3** during SEC (Figure 4.6A red chromatogram). Upon the addition of 10% PEG-3350 into a protein sample containing freshly prepared **3** at 0.5 mg/mL, a substantial peak corresponding to **1′** was observed in the SEC chromatogram within 4 hours, although no capsid was detected after the first hour (Figure 4.7). Prolonging the assembly time from 4 hours up to 96 hours resulted in a further increase in percent yield of capsid from 40% to 86%. This final yield of AaLS-switch-pH capsid is similar to that of AaLS-switch-ox.

Both capsid assembly processes also show a similar dramatic decrease in the rate of capsid formation after the first few hours, which may reflect a steep concentration dependence and would be expected for the formation of a single particle from 12 building blocks. The major difference between these two timecourses is that AaLS-switch-pH shows a 1 hour lag during which no capsid is produced (Figure 4.7), while no lag was apparent for AaLS-switch-ox, which gave a 60% yield over the first hour (Figure 3.7). While it is not clear what the different behaviors at early time points (0 to 1 hour) means for the assembly mechanisms of these two systems, it is worth noting that lag phases have been observed for virus capsid assembly *in vitro*. Such lags have been ascribed to the formation of assembly nuclei and intermediates.[38]

Another important difference in the assembly behavior of AaLS-switch-pH and AaLS-switch-ox is that no intermediate is apparent in the former system, while an intermediate was observed to accumulate in the latter system (Figure 3.7 versus Figure 4.7). This dissimilarity points to distinct assembly mechanisms for these two proteins. A more detailed understanding of the assembly mechanism for AaLS-switch-pH will require

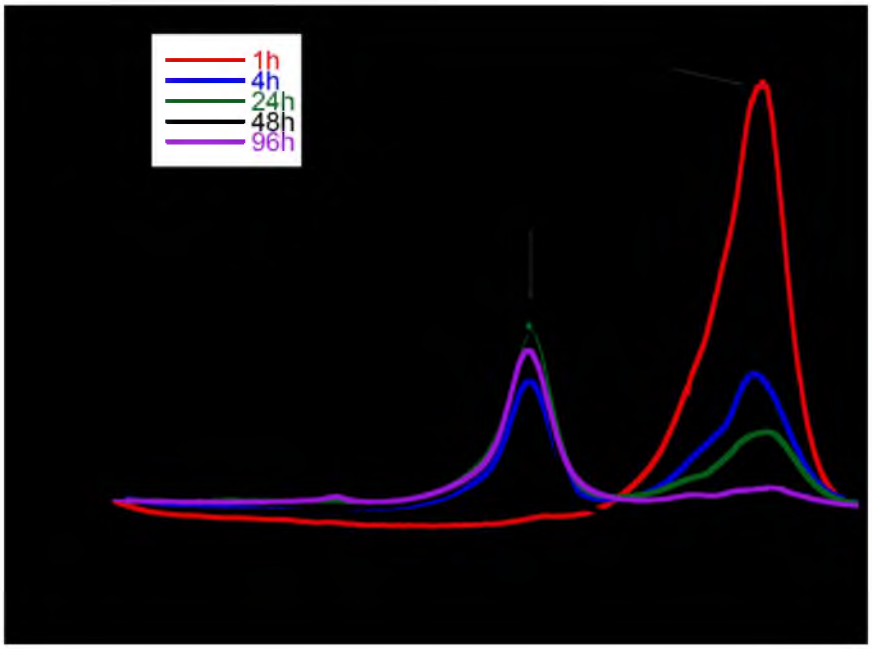


Figure 4.7. Timecourse study of AaLS-switch-pH capsid assembly at pH 8.0. Size-exclusion chromatograms for protein samples injected at 1 h (red), 4 h (blue), 24 h (green), 48 h (black), and 96 h (purple) after the addition of PEG-3350 to **3**. The total protein concentration at the start of the assembly reaction was 0.5 mg/mL. The sample was incubated at room temperature and analyzed by SEC on a HiPrep 16/60 Sephacryl S-400 HR column at 4 °C.

further study.

While PEG-3350 is essential for efficient assembly of AaLS-switch-pH capsids *in vitro,* it is not necessary for capsid stability. Indeed, after removing the PEG-3350 by ultrafiltration and SEC, **1′** showed no significant dissociation over 3 weeks at 4 °C (Figure 4.6B). As with AaLS-switch-ox, PEG-3350 likely improves the kinetics of AaLS-switch-pH assembly by facilitating the association of subunits via macromolecular crowding effects.

4.2.4. The Dependence of Capsid Disassembly on Buffer and pH

To further study the pH-dependent disassembly behavior of AaLS-switch-pH inbetween pH 8.0 and pH 5.7, **1′** was dialyzed into buffers containing 50 mM sodium phosphate and 200 mM NaCl at different pH values and then analyzed by SEC (Figure 4.8). The capsid peak was found to diminish in size as the pH decreased from 8.0 to 6.2 and the later eluting species became more prevalent; however, complete dissociation was not achieved at pH 6.2. The exact assembly states of the later eluting species in this pH range are not clear from the chromatograms. At pH values below 6.2, precipitation was observed in phosphate buffer; therefore, 50 mM sodium citrate and 200 mM NaCl was used as the buffer at pH values of 6.2 and below. In citrate buffer at pH 6.2, no capsid peak was observed, which implies that both pH and buffer identity contribute to this engineered capsid assembly switch. In addition, the bulk of the transition from capsid to pentamers occurs between pH 7.3 and pH 6.2, which is consistent with the idea that protonation of the engineered histidines drives capsid disassembly via charge repulsion.

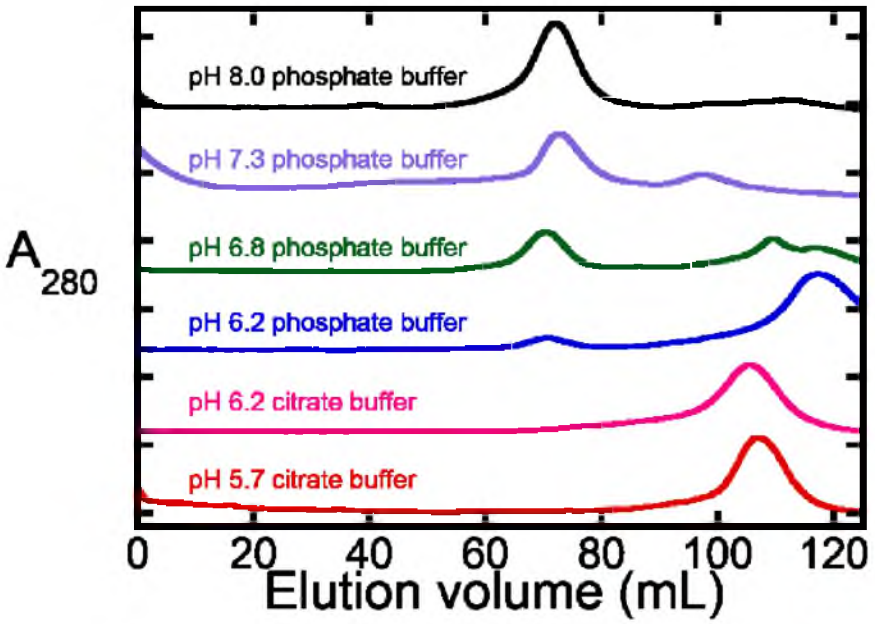


Figure 4.8. Assembly states of AaLS-switch-pH in different buffer conditions. SEC chromatograms are shown for **1′** (black), as well as samples of **1′** that had been further dialyzed into buffer containing either 50 mM sodium phosphate and 200 mM NaCl at pH values of 7.3 (purple), 6.8 (green), and 6.2 (blue), or 50 mM sodium citrate and 200 mM NaCl at pH values of 6.2 (pink), and 5.7 (red). The last sample listed corresponds to **2′**.

4.2.5. Secondary Structure and Thermostability of AaLS-switch-pH

Circular dichroism (CD) spectroscopic studies were performed to investigate the consequences of the mutations and quaternary structure changes for secondary structure and stability (Figure 4.9). While the far-UV CD spectrum of **1** is consistent with an overall α/β fold similar to wild-type AaLS (based on the SCOPe database),[155,156] slight changes in both shape and intensity were apparent. Interestingly, a much more dramatic change was found in the spectrum of **2**, indicating a large scale structural reordering relative to **1**. Indeed, the low ellipticity value at 208 nm found in the far-UV spectrum of **2** likely reflects the complete loss of α-helix at pH 5.7, and the overall shape is consistent with a predominantly β-sheet fold. In contrast, the spectrum of **3** closely resembles that of **1**, which means that the secondary structure rearrangement in the low pH environment is fully reversible upon raising the pH back from 5.7 to 8.0. However, the secondary structure changes are not accompanied by any change in quaternary structure in the transition of **1** to **3**. The spectra of **1** and **1′** suggest that the reassembled capsid has similar secondary structure to the original capsid (Figure 4.9A).

The relative stabilities of the different forms of AaLS-switch-pH were studied by measuring melting curves (Figure 4.9B). Both capsid forms, **1** and **1′**, showed similar melting behaviors; neither thermal unfolding transition began before 80 °C and neither completely unfolded by 99 °C. While the AaLS-switch-pH capsid retains a great deal of thermostability, the mutations do destabilize the capsid somewhat relative to wild-type AaLS, which showed no significant change up to 99 °C. Interestingly, the pentameric forms of AaLS-switch-pH (**2** and **3**) are less stable than the capsid form and also display significant differences in thermostability at different pH values. The unfolding of **3**,

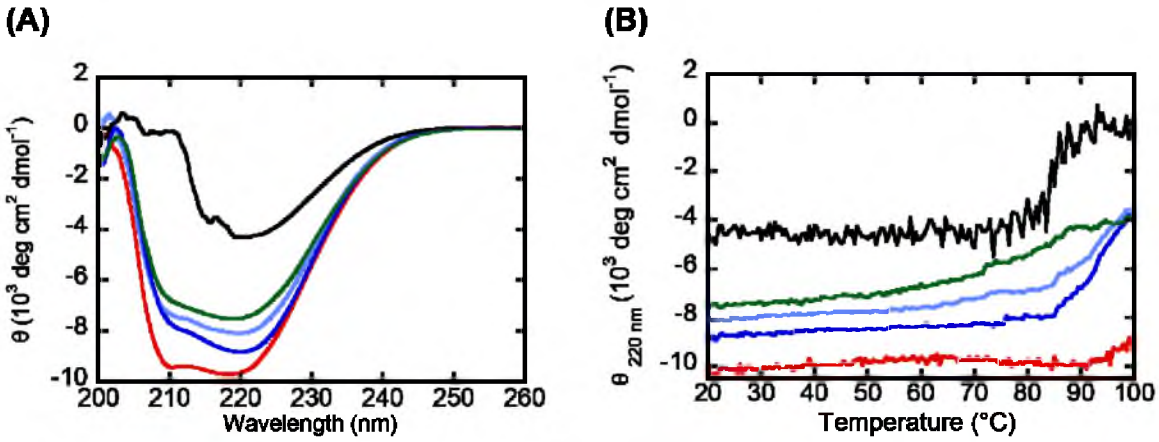


Figure 4.9. Protein secondary structure and stability analysis by CD spectroscopy. (A) Far-UV spectra were recorded in buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0 for wild-type AaLS (red), **1** (light blue), **3** (green), and **1′** (dark blue). A far-UV spectrum was also obtained in buffer containing 5 mM sodium citrate and 200 mM NaCl at pH 5.7 for **2** (black). All spectra were recorded at 20 °C. (B) Melting curves were recorded in buffer containing 50 mM sodium phosphate and 200 mM NaCl at pH 8.0 for wild-type AaLS (red), **1** (light blue), **3** (green), and **1′** (dark blue). A melting curve for **2** was recorded in buffer containing 5 mM sodium citrate and 200 mM NaCl at pH 5.7 (black). The protein concentrations were 5 μM (monomer basis), and the cuvette pathlength was 1 cm in all cases.

which was carried out in the same buffer as **1** and **1′**, began at a lower temperature and was less cooperative, although the unfolding was incomplete at 99 °C. The thermal unfolding behavior of **2** at pH 5.7 did show apparent two state unfolding with the midpoint of the transition occurring at ~ 84 °C. The different shapes of the thermal unfolding curves for **2** and **3** further highlight the pH-dependent structural reorganization observed for the pentameric AaLS-switch-pH at pH 8.0 and 5.7 (Figure 4.9B). While none of the observed thermal unfolding transitions were reversible, every form of AaLS-switch-pH remains largely folded at temperatures up to or above 60 °C.

4.3. Conclusions

In large symmetric protein complexes, a small number of mutations can have dramatic consequences for assembly.[141,142,157] Such mutations can be particularly powerful when targeted to symmetry axes. The design of AaLS-switch-pH focused on the two-fold and three-fold symmetric protein-protein interfaces, which provide important contributions to AaLS capsid stability. The interconversion of assembly states in AaLS-swtich-pH relies in part upon mutations (R40S and H41S) that remove an electrostatic interaction and a hydrogen bonding interaction at the two-fold symmetry axis. In addition, it was also necessary to introduce three histidines per subunit adjacent to the three-fold symmetric, interfacial hydrophobic cluster (Figure 4.1). The newly introduced histidines are solvent-exposed and lie at the interior surface of the capsid. At pH 8.0, these mutations minimally destabilize the capsid because histidine is polar but uncharged. However, upon lowering the pH, these residues become increasingly protonated, which generates charge repulsion. Thus, in the context of the R40S and the H41S mutations, the balance between assembled

capsids and free pentamers is determined by a trade-off between the native capsid-stabilizing hydrophobic cluster and the engineered capsid-destabilizing charge repulsion.

While charge repulsion undoubtedly causes capsid dissociation, it may do so indirectly. In wild-type AaLS, the closest approach across the pentamer-pentamer interface of the residues targeted for replacement with histidine is about 6 Å. However, within the same polypeptide, charge-charge interactions between the engineered histidines are likely to be more pronounced. Indeed, the side chains of T120 and Q123 in the wild-type AaLS structure are separated by a single turn of an α-helix (~4 Å). Unfavorable intramolecular interactions between the side chains could drive backbone rearrangements that cause capsid disassembly by disrupting intermolecular packing. This model is supported by the profound change observed in the far-UV CD spectrum of **2** relative to **1**. Further studies will be required to elucidate the individual contributions of the engineered histidines.

Many virus capsids also switch their quaternary structures in response to changes in the environmental pH value.[82,87] Often the pH-titratable residues responsible for controlling assembly states have not been identified, although structural and biochemical studies have provided insights into the acid induced disassembly of foot-and-mouth disease virus (FMDV) capsid.[158] In FMDV, four different proteins form 12 pentagonal building blocks that arrange into an icosahedral capsid. Interestingly, the disassembly of the FMDV capsid occurs over a similar pH range as AaLS-switch-pH and is also thought to rely on histidine protonation. In FMDV, the key histidine is located at the two-fold symmetric pentamer-pentamer interface. The crystal structure suggests that protonation of this histidine could lead to charge repulsion with an α-helix dipole and two other

histidines on a neighboring pentamer. In line with this expectation, replacement of the key histidine with arginine led to decreased capsid assembly while the histidine to aspartic acid variant gave acid-stable capsids. Thus, the placement of histidine near the interfaces between building blocks could provide a general strategy for developing pH-dependent capsid assembly switches both in nature and the laboratory.

As discussed in Chapter 1, engineered assembly switches have also been reported for ferritin, pyruvate dehydrogenase, and CCMV capsids.[88–90] In these cases, as in AaLS-switch-pH, the interfaces within lower order oligomeric building blocks are left intact while those between the building blocks are targeted for rational modification. The pyruvate dehydrogenase switch is the most similar to AaLS-switch-pH because both use pH titration of histidines near the three-fold symmetry axis to trigger the disassembly of dodecahedral capsids.[90] However, unlike AaLS-switch-pH, the crucial histidines in the pyruvate dehydrogenase switch are also present in the wild-type protein and the protein-protein interface was destabilized by a large N-terminal truncation. The CCMV switch also relies on pH changes and engineered histidines to promote capsid assembly, but the role of the histidines in this switch is to stabilize the capsid via metal coordination.[89] The ferritin capsid assembly switch involves the destabilization of one part of the subunit interface and the addition of a metal coordination site that bridges another part of the subunit interface.[88] In this case, metal ion titration is used to control capsid assembly by altering the balance between capsid stabilizing and destabilizing interactions, which is reminiscent of the role played by pH-titration in AaLS-switch-pH.

While the previously reported capsid assembly switches discussed above utilize a variety of protein scaffolds and design strategies, most of them work in only one direction.

The ability to assemble and disassemble reversibly is an important feature of AaLS-switch-pH. The greater control over capsid assembly offered by AaLS-switch-pH should be advantageous in the development of this scaffold for bionanotechnology.

The AaLS capsid has demonstrated promising results as a platform for developing novel encapsulation systems. However, guest loading of AaLS capsids has so far been limited to *in vivo* coproduction or *in vitro* mixing with highly cationic cargoes.[111,112,114] The ability to control the assembly and disassembly of this scaffold should enable new modes of guest loading for AaLS-based encapsulation systems suitable for a wider variety of cargo molecules. Knowledge of important interactions at the pentamer-pentamer interfaces in the AaLS capsid was exploited in the design of both AaLS-switch-ox and AaLS-switch-pH. Interactions at the three-fold symmetric interface of the capsid were essential for these switches. In AaLS-switch-ox, a stabilizing interaction was removed by mutagenesis and restored by covalent modification. The presence or absence of a prosthetic group is used to control the assembly state. In AaLS-switch-pH, the major native stabilizing interactions were left intact and new destabilizing interactions were introduced. To change the assembly state, the competition between these opposing interactions is modulated by changing the pH. While in principle both switches are reversible, in practice cycling between assembly states has been easier to accomplish with AaLS-switch-pH. Thus, this pH-dependent capsid assembly switch should increase the versatility of the AaLS scaffold for bionanotechnological applications such as engineering novel drug delivery carriers or nanoreactors.

4.4. Materials and Methods

4.4.1. Materials

Chemicals, enzymes, cell culture media, and salts were obtained from the following suppliers: Sigma-Aldrich, Thermo Fisher scientific, Macron chemicals, Gold Biotechnology, and MO-BIO laboratories (unless otherwise mentioned). All purchased chemicals were used without further purification in this study. *Pfu*-turbo DNA polymerase was obtained from Agilent Technologies. *E. coli* strains BL21 (DE3) and XL1-Blue were purchased from Stratagene. DpnI endonuclease was purchased from New England Biolabs. The oligonucleotides used in this study were synthesized by the DNA/peptide synthesis Core Facility at the University of Utah. UV-vis absorbance data were recorded using a Nanodrop 1000 spectrophotometer.

4.4.2. DNA Mutagenesis

AaLS-switch-pH is a variant of AaLS containing five point mutations (R40S, H41S, T120H, E122H, and Q123H) per polypeptide. AaLS-switch-pH was generated by performing two successive rounds of site-directed mutagenesis PCR according to the procedure described in Chapter 2. Plasmid pMG-AaLSNoHis-R40S/H41S/I125S (which encodes R40S/H41S/I125S-AaLS, as described in Chapter 2) was used as the DNA template. The sequences of the mutagenic primers are listed in Table 4.2. The resulting plasmid, pMG-AaLSNoHis-switch-pH, encodes AaLS-switch-pH (R40S/H41S/T120H/E122H/Q123H-AaLS) and was confirmed by DNA sequencing (University of Utah DNA sequencing core facility).

Table 4.2. List of mutagenic oligonucleotides used in site-directed mutagenesis PCR. Mutagenic bases are in bold.

Primer name	DNA sequence (5′ → 3′)
HCN131	GTTATTACAGCTGAC**CAT**TTGGAACA**T**GCTATCGAGCGCGCCG
HCN132	CGGCGCGCTCGATAGC**A**TGTTCCAA**ATG**GTCAGCTGTAATAAC
HCN137	CAGCTGACCATTTG**C**A**C**CATGCTATCGAGCGCGC
HCN138	GCGCGCTCGATAGCATG**G**T**G**CAAATGGTCAGCTG

4.4.3. Protein Production and Purification

AaLS-switch-pH was produced using the procedure outlined in Chapter 3. The plasmid pMG-AaLSNoHis-switch-pH was transformed into $CaCl_2$-competent *E.coli* BL21 (DE3) cells, which were cultured in LB medium (0.5 L) containing ampicillin (100 mg/L) at 37 °C and 250 rpm until the O.D. $_{600\ nm}$ reached ~0.7. Protein production was induced by the addition of IPTG (0.1 mM final concentration). Following an additional 20 h incubation at 30 °C and 250 rpm, the cultures were spun at 6000 x *g* for 10 min at 4 °C. The resulting cell pellets were stored at -20 °C until use.

AaLS-switch-pH was purified essentially as described in Chapter 3. Briefly, a cell pellet was resuspended in lysis buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8.0) containing lysozyme (5 mg), DNase (10 μg), RNase A (0.6 mg), and $MgCl_2$ (10 mM) followed by sonication. The cell lysate was incubated in a water bath at 60 °C for 5 min. The cell debris and precipitated protein were removed by centrifugation at 8000 x *g* for 30 min at 4 °C. The cleared cell lysate was dialyzed into ion-exchange buffer A (50 mM sodium phosphate, pH 8.0) overnight at 4°C. The next day, an ÄKTA FPLC system (GE HealthCare) was used to load the cell lysate onto a MonoQ 5/50 GL anion-exchange column, which had been preequilibrated with ion-exchange buffer A. AaLS-switch-pH was eluted by increasing the concentration of ion-exchange buffer B (50 mM sodium phosphate, 1 M NaCl, pH 8.0) from 5–70% over 70 mL. Fractions containing AaLS-switch-pH were pooled, concentrated, and further purified over a HiPrep 16/60 Sephacryl S400 HR SEC column that had been preequilibrated with SEC running buffer (50 mM sodium phosphate, 200 mM NaCl, pH 8.0) at a flow rate of 0.5 mL/min and 4 °C. Protein concentrations were determined using The Coomassie Plus (Bradford) assay reagent

(Thermo Scientific) according to the manufacturer's manual. Typically, 30–40 mg of purified AaLS-switch-pH can be obtained per L of cell culture.

The production and purification of wild-type AaLS and R40S/H41S/I125S-AaLS from the plasmids pMG-AaLSNoHis and pMG-AaLSNoHis-R40S/H41S/I125S, respectively, were produced and purified as previously reported in Chapter 2.

4.4.4. Assembly and Disassembly of AaLS-switch-pH Capsids *In Vitro*

To disassemble the AaLS-switch-pH capsid, freshly purified protein, **1,** was dialyzed in pH 5.7 buffer (containing 50 mM sodium citrate, 200 mM NaCl, and 5 mM EDTA) overnight at 4 °C. Following SEC, as described below (in section 4.4.5), the fractions containing disassembled AaLS-switch-pH protein, **2**, were collected and used for further characterization or for reassembly of the AaLS-switch-pH capsid. To reassemble the capsid, **2** was first dialyzed into pH 8.0 buffer (containing 50 mM sodium phosphate, 200 mM NaCl, and 5 mM EDTA) overnight at 4 °C, giving **3**, followed by the addition of 25 % (w/v) PEG-3350 stock solution into the protein solution to a final concentration of 10% (w/v) PEG-3350. The mixture of protein and PEG-3350 was stirred for at least 2 days at room temperature. After 2 days, the PEG-3350 was removed by buffer exchange with the same pH 8.0 buffer lacking PEG-3350 in a Vivaspin Protein concentrator (Sartorius) at 8000 x *g*, followed by SEC (as described below in section 4.4.5). Fractions containing the reassembled AaLS-switch-pH capsid, **1′**, were pooled, concentrated, and used either for the characterization or for a second round of disassembly, producing **2′**, using the same procedure described above for the first disassembly at pH 5.7. The

assembly state of **2′** was analyzed by SEC as described in the next section.

4.4.5. Analysis of Protein Assembly State by Size-exclusion Chromatography

The assembly states of AaLS-switch-pH following different buffer treatments were analyzed by SEC on an ÄKTA FPLC system (GE Healthcare). In brief, 1 mL of protein sample (0.4–1.0 mg/mL) was injected onto a HiPrep 16/16 Sephacryl S-400 HR column which bad been preequilibrated with pH 8.0 running buffer (50 mM sodium phosphate and 200 mM NaCl) for **1**, **3**, and **1′**, or pH 5.7 running buffer (50 mM sodium citrate and 200 mM NaCl) for **2** and **2′**. The flow rate was 0.5 mL/min and the running temperature was 4 °C. The assembly states of AaLS-switch-pH were determined based on comparison to wild-type AaLS (1100 kDa, 60-subunit capsid)[103] and R40S/H41S/I125S-AaLS (84 kDa, pentamer, described in Chapter 2).

4.4.6. Sedimentation Equilibrium Analysis

The samples of **1**, **2**, and **1′** were analyzed by sedimentation equilibrium in a Beckman XL-A analytical ultracentrifuge equipped with An60Ti rotor and photoelectric scanner using the protocol reported in Chapter 2. Species **1** and **1′** were spun at 4000 rpm and 5000 rpm and at three different protein concentrations (0.5 mg/mL, 0.75 mg/mL, and 1 mg/mL) in pH 8.0 buffer (50 mM sodium phosphate and 200 mM NaCl). Species **2** was spun at 12500, 15000, and 20000 rpm and three different protein concentrations, as listed above in pH 5.7 buffer (50mM sodium citrate and 200 mM NaCl). The data were analyzed using Ultrascan II software (University of Texas Health Science Center) and fit to a single

ideal species model, as described in Chapter 2.

4.4.7. Circular Dichroism Spectroscopy

The CD spectroscopy was performed according the method reported in Chapter 2. The far-UV CD spectra and thermal denaturation curves of wild-type AaLS, **1**, **1′**, and **3** were recorded in pH 8.0 buffer (50 mM sodium phosphate and 200 mM NaCl). The CD spectrum and thermal denaturation curve of **2** were recorded in pH 5.7 buffer (5 mM sodium citrate and 200 mM NaCl). The low sodium citrate concentration was required to minimize background signal generated by the buffer. All CD spectra and thermal denaturation curves were obtained using 5 μM protein solutions in a 1-cm pathlength cuvette.

4.4.8. Transmission Electron Microscopy

Protein samples of **1** and **1′** were imaged by negative-stain TEM at 125 kV accelerating voltage. Samples were prepared using the protocol described in Chapter 3. Each sample consisted of AaLS-switch-pH (0.1 mg/mL) in pH 8.0 buffer (containing 50 mM sodium phosphate and 200 mM NaCl). Wild-type AaLS capsid was also visualized the same way as a control. TEM images were recorded using a Hitachi H7100 transmission electron microscope and analyzed by using ImageJ (National Institutes of Health). The particle images used in the analysis were selected by limiting the circularity values to 0.15–1.0 and an area of 100–500 nm^2. A total of 127 particles and 154 particles were picked analyzed for **1** and **1′**, respectively.

Averaged TEM images for wild-type AaLS, **1** and **1′** were generated using the

program Bsoft.[159] For wild-type AaLS, **1**, and **1′**, 46, 31, and 35 individual capsid particles were selected to generate the averaged images, respectively. The averaged images were analyzed using ImageJ (National Institutes of Health). The gray value was recorded across four different cross sections (representing 90° rotations) of each capsid to generate plots of the gray value versus the distance across the averaged image.

4.4.9. Energy Minimization of The Three-fold Symmetry Axis of AaLS-switch-pH

A model of the three-fold symmetry axis of AaLS-switch-pH was constructed from a fragment of the crystal structure of wild-type AaLS (PDB id: 1HQK) consisting of three adjacent pentamers. Residues T120, E122, and Q123 of wild-type AaLS were mutated to histidine using the program SYBYL-X 2.0 (Tripos). This model of the T120H/E122H/Q123H-AaLS variant was energy minimized using the Powell method and AMBER7 FF02 force field over the region within 8 Å of the mutated histidines. The default parameters were selected with the following exceptions: Charges (No charge), Dielectric Function (Distance), Termination (Gradient, 0.05 kcal/(mol·A)), and Maximum Iterations (1,000,000). The resulting structure was visualized using Pymol (DeLano Scientific LLC).

REFERENCES

(1) Goodsell, D. S.; Olson, A. J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105.

(2) Janin, J.; Bahadur, R. P.; Chakrabarti, P. *Q. Rev. Biophys.* **2008**, *41*, 133.

(3) Walden, H.; Bell, G. S.; Russell, R. J.; Siebers, B.; Hensel, R.; Taylor, G. L. *J. Mol. Biol.* **2001**, *306*, 745.

(4) Crick, F. H.; Watson, J. D. *Nature* **1956**, *177*, 473.

(5) Kurland, C. G. *Annu. Rev. Genet.* **1992**, *26*, 29.

(6) Parker, J. *Microbiol. Rev.* **1989**, *53*, 273.

(7) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13.

(8) Levy, E. D.; Pereira-Leal, J. B.; Chothia, C.; Teichmann, S. A. *PLoS Comput. Biol.* **2006**, *2*, e155.

(9) Venkatakrishnan, A. J.; Levy, E. D.; Teichmann, S. A. *Biochem. Soc. Trans.* **2010**, *38*, 879.

(10) Perica, T.; Marsh, J. A.; Sousa, F. L.; Natan, E.; Colwell, L. J.; Ahnert, S. E.; Teichmann, S. A. *Biochem. Soc. Trans.* **2012**, *40*, 475.

(11) Cornish-Bowden, A. J.; Koshland, D. E. *J. Biol. Chem.* **1971**, *246*, 3092.

(12) André, I.; Strauss, C. E.; Kaplan, D. B.; Bradley, P.; Baker, D. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16148.

(13) Ali, M. H.; Imperiali, B. *Bioorg. Med. Chem.* **2005**, *13*, 5013.

(14) Kuriyan, J.; Eisenberg, D. *Nature* **2007**, *450*, 983.

(15) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88.

(16) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. *Chem. Rev.* **2008**, *108*, 1225.

(17) Keskin, O.; Tuncbag, N.; Gursoy, A. *Curr. Pharm. Biotechnol.* **2008**, *9*, 67.

(18) Reichmann, D.; Rahat, O.; Cohen, M.; Neuvirth, H.; Schreiber, G. *Curr. Opin. Struct. Biol.* **2007**, *17*, 67.(19)Nooren, I. M.; Thornton, J. M. *EMBO J.* **2003**, *22*, 3486.

(20) Nooren, I. M.; Thornton, J. M. *J. Mol. Biol.* **2003**, *325*, 991.

(21) Perica, T.; Chothia, C.; Teichmann, S. A. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 8127.

(22) Jones, S.; Thornton, J. M. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 31.

(23) Bahadur, R. P.; Chakrabarti, P.; Rodier, F.; Janin, J. *Proteins* **2003**, *53*, 708.

(24) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155.

(25) Clackson, T.; Wells, J. A. *Science* **1995**, *267*, 383.

(26) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1.

(27) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. *Proteins* **2007**, *68*, 803.

(28) Yeates, T. O.; Padilla, J. E. *Curr. Opin. Struct. Biol.* **2002**, *12*, 464.

(29) Tsai, C. J.; Zheng, J.; Zanuy, D.; Haspel, N.; Wolfson, H.; Alemán, C.; Nussinov, R. *Proteins* **2007**, *68*, 1.

(30) Woolfson, D. N.; Mahmoud, Z. N. *Chem. Soc. Rev.* **2010**, *39*, 3464.

(31) Caspar, D. L.; Klug, A. *Cold Spring Harb. Symp. Quant. Biol.* **1962**, *27*, 1.

(32) Yeates, T. O.; Thompson, M. C.; Bobik, T. A. *Curr. Opin. Struct. Biol.* **2011**, *21*, 223.

(33) Aevarsson, A.; Seger, K.; Turley, S.; Sokatch, J. R.; Hol, W. G. *Nat. Struct. Biol.* **1999**, *6*, 785.

(34) Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolásek, B.; Ban, N. *Science* **2007**, *316*, 254.

(35) Yeates, T. O.; Crowley, C. S.; Tanaka, S. *Annu. Rev. Biophys.* **2010**, *39*, 185.

(36) Johnson, J. E.; Speir, J. A. *J. Mol. Biol.* **1997**, *269*, 665.

(37) Rossmann, M. G.; Rao, V. B.; Prasad, B. V. V.; Schmid, M. In *Viral Molecular Machines*; Springer US: New York, 2012; Vol. 726, p 17.

(38) Katen, S.; Zlotnick, A. *Methods Enzymol.* **2009**, *455*, 395.

(39) Zlotnick, A. *J. Mol. Recognit.* **2005**, *18*, 479.

(40) Zandi, R.; van der Schoot, P.; Reguera, D.; Kegel, W.; Reiss, H. *Biophys. J.* **2006**, *90*, 1939.

(41) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. *Virology* **2000**, *277*, 450.

(42) Zhou, Z. H.; McCarthy, D. B.; O'Connor, C. M.; Reed, L. J.; Stoops, J. K. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14802.

(43) Schott, K.; Ladenstein, R.; Konig, A.; Bacher, A. *J. Biol. Chem.* **1990**, *265*, 12686.

(44) Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. *Annu. Rev. Microbiol.* **2010**, *64*, 391.

(45) Izard, T.; Aevarsson, A.; Allen, M. D.; Westphal, A. H.; Perham, R. N.; de Kok, A.; Hol, W. G. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1240.

(46) Kis, K.; Bacher, A. *J. Biol. Chem.* **1995**, *270*, 16788.

(47) Yeates, T. O.; Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C.; Shively, J. M. *Nat. Rev. Microbiol.* **2008**, *6*, 681.

(48) Cannon, G. C.; Heinhorst, S.; Kerfeld, C. A. *Biochim. Biophys. Acta.* **2010**, *1804*, 382.

(49) Tanaka, S.; Kerfeld, C. A.; Sawaya, M. R.; Cai, F.; Heinhorst, S.; Cannon, G. C.; Yeates, T. O. *Science* **2008**, *319*, 1083.

(50) Klein, M. G.; Zwart, P.; Bagby, S. C.; Cai, F.; Chisholm, S. W.; Heinhorst, S.; Cannon, G. C.; Kerfeld, C. A. *J. Mol. Biol.* **2009**, *392*, 319.

(51) Bobik, T. A.; Havemann, G. D.; Busch, R. J.; Williams, D. S.; Aldrich, H. C. *J. Bacteriol.* **1999**, *181*, 5967.

(52) Penrod, J. T.; Roth, J. R. *J. Bacteriol.* **2006**, *188*, 2865.

(53) Sutter, M.; Boehringer, D.; Gutmann, S.; Gunther, S.; Prangishvili, D.; Loessner, M. J.; Stetter, K. O.; Weber-Ban, E.; Ban, N. *Nat. Struct. Mol. Biol.* **2008**, *15*, 939.

(54) Chasteen, N. D.; Harrison, P. M. *J. Struct. Biol.* **1999**, *126*, 182.

(55) Hempstead, P. D.; Yewdall, S. J.; Fernie, A. R.; Lawson, D. M.; Artymiuk, P. J.; Rice, D. W.; Ford, G. C.; Harrison, P. M. *J. Mol. Biol.* **1997**, *268*, 424.

(56) Harrison, P. M.; Hoy, T. G.; Macara, I. G.; Hoare, R. J. *Biochem. J.* **1974**, *143*, 445.

(57) Watt, R. K. *Biometals* **2011**, *24*, 489.

(58) Hintze, K. J.; Theil, E. C. *Cell. Mol. Life. Sci.* **2006**, *63*, 591.

(59) Watt, R. K.; Hilton, R. J.; Graff, D. M. *Biochim. Biophys. Acta.* **2010**, *1800*, 745.

(60) Uchida, M.; Kang, S.; Reichhardt, C.; Harlen, K.; Douglas, T. *Biochim. Biophys. Acta.* **2010**, *1800*, 834.

(61) Uchida, M.; Klem, M. T.; Allen, M.; Suci, P.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Liepold, L. O.; Young, M.; Douglas, T. *Adv. Mater.* **2007**, *19*, 1025.

(62) Kramer, R. M.; Li, C.; Carter, D. C.; Stone, M. O.; Naik, R. R. *J. Am. Chem. Soc.* **2004**, *126*, 13282.

(63) Kasyutich, O.; Ilari, A.; Fiorillo, A.; Tatchev, D.; Hoell, A.; Ceci, P. *J. Am. Chem. Soc.* **2010**, *132*, 3621.

(64) Douglas, T.; Young, M. *Nature* **1998**, *393*, 152.

(65) Douglas, T.; Strable, E.; Willits, D.; Aitouchen, A.; Libera, M.; Young, M. *Adv. Mater.* **2002**, *14*, 415.

(66) Abedin, M. J.; Liepold, L.; Suci, P.; Young, M.; Douglas, T. *J. Am. Chem. Soc.* **2009**, *131*, 4346.

(67) Ma, Y.; Nolte, R. J.; Cornelissen, J. J. *Adv. Drug Deliv. Rev.* **2012**, *64*, 811.

(68) Mateu, M. G. *Protein Eng. Des. Sel.* **2011**, *24*, 53.

(69) Kovacs, E. W.; Hooker, J. M.; Romanini, D. W.; Holder, P. G.; Berry, K. E.; Francis, M. B. *Bioconjug. Chem.* **2007**, *18*, 1140.

(70) Datta, A.; Hooker, J. M.; Botta, M.; Francis, M. B.; Aime, S.; Raymond, K. N. *J. Am. Chem. Soc.* **2008**, *130*, 2546.

(71) Farkas, M. E.; Aanei, I. L.; Behrens, C. R.; Tong, G. J.; Murphy, S. T.; O'Neil, J. P.; Francis, M. B. *Mol. Pharm.* **2013**, *10*, 69.

(72) Liepold, L.; Anderson, S.; Willits, D.; Oltrogge, L.; Frank, J. A.; Douglas, T.; Young, M. *Magn. Reson. Med.* **2007**, *58*, 871.

(73) Yamada, T.; Iwasaki, Y.; Tada, H.; Iwabuki, H.; Chuah, M. K.; VandenDriessche, T.; Fukuda, H.; Kondo, A.; Ueda, M.; Seno, M.; Tanizawa, K.; Kuroda, S. *Nat. Biotechnol.* **2003**, *21*, 885.

(74) Ren, Y.; Wong, S. M.; Lim, L. Y. *Bioconjug. Chem.* **2007**, *18*, 836.

(75) Kwak, M.; Minten, I. J.; Anaya, D. M.; Musser, A. J.; Brasch, M.; Nolte, R. J.; Müllen, K.; Cornelissen, J. J.; Herrmann, A. *J. Am. Chem. Soc.* **2010**, *132*, 7834.

(76) Bode, S. A.; Minten, I. J.; Nolte, R. J.; Cornelissen, J. J. *Nanoscale* **2011**, *3*, 2376.

(77) Minten, I. J.; Claessen, V. I.; Blank, K.; Rowan, A. E.; Nolte, R. J. M.; Cornelissen, J. J. L. M. *Chemical Science* **2011**, *2*, 358.

(78) Patterson, D. P.; Prevelige, P. E.; Douglas, T. *ACS Nano* **2012**, *6*, 5000.

(79) Fiedler, J. D.; Brown, S. D.; Lau, J. L.; Finn, M. G. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 9648.

(80) Comellas-Aragones, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N. A.; Rowan, A. E.; Christianen, P. C.; Maan, J. C.; Verduin, B. J.; Cornelissen, J. J.; Nolte, R. J. *Nat. Nanotechnol.* **2007**, *2*, 635.

(81) Speir, J. A.; Munshi, S.; Wang, G.; Baker, T. S.; Johnson, J. E. *Structure* **1995**, *3*, 63.

(82) Lavelle, L.; Michel, J. P.; Gingery, M. *J. Virol. Methods.* **2007**, *146*, 311.

(83) Minten, I. J.; Hendriks, L. J.; Nolte, R. J.; Cornelissen, J. J. *J. Am. Chem. Soc.* **2009**, *131*, 17771.

(84) O'Neil, A.; Reichhardt, C.; Johnson, B.; Prevelige, P. E.; Douglas, T. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 7425.

(85) Ceres, P.; Zlotnick, A. *Biochemistry* **2002**, *41*, 11525.

(86) Singh, S.; Zlotnick, A. *J. Biol. Chem.* **2003**, *278*, 18249.

(87) Chuan, Y. P.; Fan, Y. Y.; Lua, L. H.; Middelberg, A. P. *J. R. Soc. Interface* **2010**, *7*, 409.

(88) Huard, D. J.; Kane, K. M.; Tezcan, F. A. *Nat. Chem. Biol.* **2013**, *9*, 169.

(89) Minten, I. J.; Wilke, K. D.; Hendriks, L. J.; van Hest, J. C.; Nolte, R. J.; Cornelissen, J. J. *Small* **2011**, *7*, 911.

(90) Dalmau, M.; Lim, S.; Wang, S. W. *Biomacromolecules* **2009**, *10*, 3199.

(91) van Eldijk, M. B.; Wang, J. C.; Minten, I. J.; Li, C.; Zlotnick, A.; Nolte, R. J.; Cornelissen, J. J.; van Hest, J. C. *J. Am. Chem. Soc.* **2012**, *134*, 18506.

(92) Edgcomb, S. P.; Murphy, K. P. *Proteins* **2002**, *49*, 1.

(93) Miyagi, M.; Nakazawa, T. *Anal. Chem.* **2008**, *80*, 6481.

(94) Urry, D. W.; Pattanaik, A. *Ann. N. Y. Acad. Sci.* **1997**, *831*, 32.

(95) van Eldijk, M. B.; McGann, C. L.; Kiick, K. L.; van Hest, J. C. *Top. Curr. Chem.* **2012**, *310*, 71.

(96) Bacher, A.; Eberhardt, S.; Eisenreich, W.; Fischer, M.; Herz, S.; Illarionov, B.; Kis, K.; Richter, G. *Vitam. Horm.* **2001**, *61*, 1.

(97) Meining, W.; Mortl, S.; Fischer, M.; Cushman, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2000**, *299*, 181.

(98) Gerhardt, S.; Haase, I.; Steinbacher, S.; Kaiser, J. T.; Cushman, M.; Bacher, A.; Huber, R.; Fischer, M. *J. Mol. Biol.* **2002**, *318*, 1317.

(99) Persson, K.; Schneider, G.; Jordan, D. B.; Viitanen, P. V.; Sandalova, T. *Protein Sci.* **1999**, *8*, 2355.

(100) Morgunova, E.; Saller, S.; Haase, I.; Cushman, M.; Bacher, A.; Fischer, M.; Ladenstein, R. *J. Biol. Chem.* **2007**, *282*, 17231.

(101) Morgunova, E.; Meining, W.; Illarionov, B.; Haase, I.; Jin, G.; Bacher, A.; Cushman, M.; Fischer, M.; Ladenstein, R. *Biochemistry* **2005**, *44*, 2746.

(102) Zylberman, V.; Craig, P. O.; Klinke, S.; Braden, B. C.; Cauerhff, A.; Goldbaum, F. A. *J. Biol. Chem.* **2004**, *279*, 8093.

(103) Zhang, X.; Meining, W.; Fischer, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2001**, *306*, 1099.

(104) Ladenstein, R.; Ritsert, K.; Huber, R.; Richter, G.; Bacher, A. *Eur. J. Biochem.* **1994**, *223*, 1007.

(105) Mortl, S.; Fischer, M.; Richter, G.; Tack, J.; Weinkauf, S.; Bacher, A. *J. Biol. Chem.* **1996**, *271*, 33201.

(106) Kumar, P.; Singh, M.; Karthikeyan, S. *Acta. Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 131.

(107) Zhang, X.; Konarev, P. V.; Petoukhov, M. V.; Svergun, D. I.; Xing, L.; Cheng, R. H.; Haase, I.; Fischer, M.; Bacher, A.; Ladenstein, R.; Meining, W. *J. Mol. Biol.* **2006**, *362*, 753.

(108) Ladenstein, R.; Fischer, M.; Bacher, A. *FEBS J.* **2013**, *280*, 2537.

(109) Braun, N.; Meining, W.; Hars, U.; Fischer, M.; Ladenstein, R.; Huber, R.; Bacher, A.; Weinkauf, S.; Bachmann, L. *J. Mol. Biol.* **2002**, *321*, 341.

(110) Shenton, W.; Mann, S.; Cölfen, H.; Bacher, A.; Fischer, M. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 442.

(111) Seebeck, F. P.; Woycechowsky, K. J.; Zhuang, W.; Rabe, J. P.; Hilvert, D. *J. Am. Chem. Soc.* **2006**, *128*, 4516.

(112) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D. *Science* **2011**, *331*, 589.

(113) Lilavivat, S.; Sardar, D.; Jana, S.; Thomas, G. C.; Woycechowsky, K. J. *J. Am. Chem. Soc.* **2012**, *134*, 13152.

(114) Wörsdörfer, B.; Pianowski, Z.; Hilvert, D. *J. Am. Chem. Soc.* **2012**, *134*, 909.

(115) Grueninger, D.; Treiber, N.; Ziegler, M. O.; Koetter, J. W.; Schulze, M. S.; Schulz, G. E. *Science* **2008**, *319*, 206.

(116) André, I.; Bradley, P.; Wang, C.; Baker, D. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17656.

(117) Padilla, J. E.; Colovos, C.; Yeates, T. O. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 2217.

(118) Douglas, T.; Young, M. *Science* **2006**, *312*, 873.

(119) Papapostolou, D.; Howorka, S. *Mol. Biosyst.* **2009**, *5*, 723.

(120) Lau, J. L.; Baksh, M. M.; Fiedler, J. D.; Brown, S. D.; Kussrow, A.; Bornhop, D. J.; Ordoukhanian, P.; Finn, M. G. *ACS Nano* **2011**, *5*, 7722.

(121) Wu, W.; Hsiao, S. C.; Carrico, Z. M.; Francis, M. B. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 9493.

(122) Stephanopoulos, N.; Tong, G. J.; Hsiao, S. C.; Francis, M. B. *ACS Nano* **2010**, *4*, 6014.

(123) Uchida, M.; Morris, D. S.; Kang, S.; Jolley, C. C.; Lucon, J.; Liepold, L. O.; LaFrance, B.; Prevelige, P. E.; Douglas, T. *Langmuir* **2012**, *28*, 1998.

(124) Aniagyei, S. E.; Kennedy, C. J.; Stein, B.; Willits, D. A.; Douglas, T.; Young, M. J.; De, M.; Rotello, V. M.; Srisathiyanarayanan, D.; Kao, C. C.; Dragnea, B. *Nano Lett.* **2009**, *9*, 393.

(125) Polanams, J.; Ray, A. D.; Watt, R. K. *Inorg. Chem.* **2005**, *44*, 3203.

(126) Lu, X.; Thompson, J. R.; Perry, K. L. *J. Gen. Virol.* **2012**, *93*, 1120.

(127) Fraenkel-Conrat, H.; Williams, R. C. *Proc. Natl. Acad. Sci. USA* **1955**, *41*, 690.

(128) Zhao, X.; Fox, J. M.; Olson, N. H.; Baker, T. S.; Young, M. J. *Virology* **1995**, *207*, 486.

(129) Ainciart, N.; Zylberman, V.; Craig, P. O.; Nygaard, D.; Bonomi, H. R.;

Cauerhff, A. A.; Goldbaum, F. A. *Proteins* **2011**, *79*, 1079.

(130) Fornasari, M. S.; Laplagne, D. A.; Frankel, N.; Cauerhff, A. A.; Goldbaum, F. A.; Echave, J. *Mol. Biol. Evol.* **2004**, *21*, 97.

(131) Braden, B. C.; Velikovsky, C. A.; Cauerhff, A. A.; Polikarpov, I.; Goldbaum, F. A. *J. Mol. Biol.* **2000**, *297*, 1031.

(132) Woycechowsky, K. J.; Seebeck, F. P.; Hilvert, D. *Protein Sci.* **2006**, *15*, 1106.

(133) Beernink, P. T.; Tolan, D. R. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 5374.

(134) Beernink, P. T.; Tolan, D. R. *Protein Sci.* **1994**, *3*, 1383.

(135) Horovitz, A.; Bochkareva, E. S.; Girshovich, A. S. *J. Biol. Chem.* **1993**, *268*, 9957.

(136) Sabio, G.; Mora, A.; Rangel, M. A.; Quesada, A.; Marcos, C. F.; Alonso, J. C.; Soler, G.; Centeno, F. *FEBS Lett.* **2001**, *501*, 161.

(137) Jones, D. H.; McMillan, A. J.; Fersht, A. R.; Winter, G. *Biochemistry* **1985**, *24*, 5852.

(138) Bahadur, R. P.; Rodier, F.; Janin, J. *J. Mol. Biol.* **2007**, *367*, 574.

(139) Velichko, I. S.; Mikalahti, K.; Kasho, V. N.; Dudarenkov, V. Y.; Hyytiä, T.; Goldman, A.; Cooperman, B. S.; Lahti, R.; Baykov, A. A. *Biochemistry* **1998**, *37*, 734.

(140) Pornillos, O.; Ganser-Pornillos, B. K.; Banumathi, S.; Hua, Y.; Yeager, M. *J. Mol. Biol.* **2010**, *401*, 985.

(141) Zhang, Y.; Raudah, S.; Teo, H.; Teo, G. W.; Fan, R.; Sun, X.; Orner, B. P. *J. Biol. Chem.* **2010**, *285*, 12078.

(142) Zhang, Y.; Fu, J.; Chee, S. Y.; Ang, E. X.; Orner, B. P. *Protein Sci.* **2011**, *20*, 1907.

(143) Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N. *Nat. Chem. Biol.* **2010**, *6*, 615.

(144) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N. *Science* **2013**, *340*, 595.

(145) King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. *Science* **2012**, *336*, 1171.

(146) Speir, J. A.; Johnson, J. E. *Curr. Opin. Struct. Biol.* **2012**, *22*, 65.

(147) Li, K.; Nguyen, H. G.; Lu, X.; Wang, Q. *Analyst* **2010**, *135*, 21.

(148) Adolph, K. W.; Butler, P. J. *Nature* **1975**, *255*, 737.

(149) Chen, H. N.; Woycechowsky, K. J. *Biochemistry* **2012**, *51*, 4704.

(150) Fu, C.-y.; Morais, M. C.; Battisti, A. J.; Rossmann, M. G.; Prevelige Jr, P. E. *J. Mol. Biol.* **2007**, *366*, 1161.

(151) Boyd, D. B. *J. Am. Chem. Soc.* **1972**, *94*, 8799.

(152) Riddles, P. W.; Blakeley, R. L.; Zerner, B. *Methods Enzymol.* **1983**, *91*, 49.

(153) Lai, Y. T.; King, N. P.; Yeates, T. O. *Trends Cell Biol.* **2012**, *22*, 653.

(154) Zlotnick, A.; Stray, S. J. *Trends Biotechnol.* **2003**, *21*, 536.

(155) Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. *J. Mol. Biol.* **1995**, *247*, 536.

(156) Fox, N. K.; Brenner, S. E.; Chandonia, J. M. *Nucleic Acids Res.* **2014**, *42*, D304.

(157) Khare, G.; Nangpal, P.; Tyagi, A. K. *Biochemistry* **2013**, *52*, 1694.

(158) Ellard, F. M.; Drew, J.; Blakemore, W. E.; Stuart, D. I.; King, A. M. *J. Gen. Virol.* **1999**, *80 (Pt 8)*, 1911.

(159) Heymann, J. B.; Belnap, D. M. *J. Struct. Biol.* **2007**, *157*, 3.